

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-116094
(43)Date of publication of application : 06.05.1998

(51)Int.Cl. G10L 3/00
G10L 3/00
(21)Application number : 09-265959 (71)Applicant : LUCENT TECHNOLOG INC
(22)Date of filing : 30.09.1997 (72)Inventor : UU CHOU
JUANG BIING HWANG
KAWAHARA TATSUYA
CHIN-FUI LEE

(30)Priority
Priority number : 96 724413 Priority date : 01.10.1996 Priority country : US
96 771732 20.12.1996 US

(54) METHOD AND DEVICE FOR VOICE RECOGNITION

(57)Abstract:
PROBLEM TO BE SOLVED: To efficiently conduct conversation voice recognition with a high degree of freedom while accepting freer conversation than the conversation based on a fixed and formalized grammar.
SOLUTION: A key phrase detector 11 detects plural key phrases based on the set of the phrase part grammar of the intrinsic condition of the conversation. Then, a key phrase verifier 12 assigns reliability scales to these key phrases, the scales are compared with threshold values to verify the key phrases and a set of verified key phrase candidates is obtained. Then, a sentence hypothesis generator 13 couples the candidates based on a prescribed (task intrinsic, for example) meaning information 25 and a sentence hypothesis 14 is obtained. Lastly, a sentence hypothesis verifier 14 verifies the sentence hypothesis, verified sentence hypothesis is generated and a final recognition result is obtained. In the conversation base system, individual phrase part grammar is used in accordance with the specific state of the conversation in a large scale task (i.e., against 'subtasks').

LEGAL STATUS

[Date of request for examination] 05.04.1999

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

(10)日本国特許庁 (J P) (12) 公開 特 許 公 報 (A)

(11)特許出願公開番号
特開平10-116094
(43)公開日 平成10年(1998)5月6日

(51)Int.Cl.* 識別記号 F I
G10L 3/00 561 G10L 3/00 561 G
531 531 C

審査請求 未請求 請求項の数32 OL (全 12 頁)

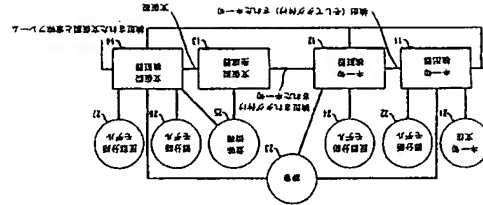
(21)出願番号 特願平9-265959
(22)出願日 平成9年(1997)9月30日
(31)優先権主張番号 08/724413
(32)優先日 1996年10月1日
(33)優先権主張国 米国 (US)
(31)優先権主張番号 08/771732
(32)優先日 1996年12月20日
(33)優先権主張国 米国 (US)
(71)出願人 59607259
ルーセント テクノロジーズ インコーポ
レイテッド
Lucent Technologies
Inc.
アメリカ合衆国 07974 ニュージャージー
ー、マレービル、マウンテン アベニュー
600-700
(72)発明者 ウー チョウ
アメリカ合衆国、07922 ニュージャージー
ー、パークレー ハイツイ、グリーンブライ
アードライブ 22
(74)代理人 弁護士 三橋 弘文
最終頁に続く

(54) [発明の名称] 音声認識方法および音声認識装置

(57) 【要約】

【課題】 固定した形式的な文法に基づいて受容されるよりも多くの発話を受容される、効率および自由度の高い会話音声認識を実現する。

【解決手段】 キー句検出器 11で、会話の状態に固有の句部分文法の集合に基づいて複数のキー句を検出する。次に、キー句検出器 12で、これらのキー句に属する性質尺度を割り当て、その信頼性尺度をしきい値と比較することによってそれらのキー句を検証し、検証済みキー句候補の集合を得る。次に、文脈生成器 13で、検証済みキー句候補を、所定の (例えばタスク固有の) 意味情報 25に基づいて結合して文脈候補を得る。最後に、文脈候補検出器 14で、これらの文脈候補を検証して、検証済み文脈候補を得る。最終認識結果を得る。会話ベースのシステムでは、大規模なタスク内の会話の特定の状態に応じて (すなわち「サブタスク」に対して)、個別の句部分文法を使用することが可能である。



【特許請求の範囲】

【請求項1】 複数の単語からなる音声発話の音声認識を実行する音声認識方法において、
部分文法に基づいてキー句抽出を実行して、認識された単語からなる複数の抽出済みキー句を生成する抽出ステップと、

前記抽出済みキー句に信頼性尺度を割り当て、該信頼性尺度を大きい値と比較することにより、前記抽出済みキー句の検証を実行して、検証済みキー句候補の集合を生成するキー句検証ステップと、
前記検証済みキー句候補を結合し、所定の意味情報に基づいて文仮説を生成するステップと、
前記文仮説の検証を実行して、少なくとも1つの検証済み文仮説を生成する文仮説検証ステップとからなることを特徴とする音声認識方法。

【請求項2】 前記部分文法は、会話状態に基づく部分文法の集合から選択されることを特徴とする請求項1の方法。

【請求項3】 前記部分文法は、音声サンプルのコーパスを用いたトレーニングプロセスに基づいて導出されたものであることを特徴とする請求項1の方法。

【請求項4】 前記文仮説の生成は、前記信頼性尺度にも基づくことを特徴とする請求項1の方法。

【請求項5】 前記抽出済みキー句は、意味タグでラベルされることを特徴とする請求項1の方法。

【請求項6】 前記文仮説の生成は、前記意味タグにも基づくことを特徴とする請求項5の方法。

【請求項7】 前記文仮説の生成は、前記信頼性尺度にも基づくことを特徴とする請求項6の方法。

【請求項8】 前記文仮説の生成は、前記信頼性尺度、前記意味タグ、および前記所定の意味情報に基づいて、最も確からしい文仮説を判定するステップからなることを特徴とする請求項7の方法。

【請求項9】 前記抽出ステップは複数の部分語モデルに基づいて実行され、前記抽出済みキー句は部分語の列からなることを特徴とする請求項1の方法。

【請求項10】 前記部分語モデルは隠れマルコフモデルからなることを特徴とする請求項9の方法。

【請求項11】 前記抽出ステップは、部分語モデルの集合と、対応する反部分語モデルの集合に基づいて実行されることを特徴とする請求項9の方法。

【請求項12】 前記部分語モデルおよび前記反部分語モデルは隠れマルコフモデルからなることを特徴とする請求項11の方法。

【請求項13】 前記文仮説検証ステップは、文仮説に対して音響的検証を実行するステップからなることを特徴とする請求項1の方法。

【請求項14】 前記文仮説検証ステップは、文仮説に対して意味的検証を実行するステップからなることを特徴とする請求項1の方法。

【請求項15】 前記文仮説検証ステップは、最も確からしい1つの文仮説を選択するステップを含むことを特徴とする請求項1の方法。

【請求項16】 前記検証済み文仮説に基づいて意味フレームを生成するステップをさらに有することを特徴とする請求項1の方法。

【請求項17】 複数の単語からなる音声発話の音声認識を実行する音声認識装置において、
句部分文法に基づいてキー句抽出を実行して、認識された単語からなる複数の抽出済みキー句を生成するキー句抽出器と、
前記抽出済みキー句候補を割り当て、該信頼性尺度を大きい値と比較することにより、前記抽出済みキー句の検証を実行して、検証済みキー句候補の集合を生成するキー句検証器と、
前記検証済みキー句候補を結合し、所定の意味情報に基づいて文仮説を生成する文仮説生成器と、
前記文仮説の検証を実行して、少なくとも1つの検証済み文仮説を生成する文仮説検証器とからなることを特徴とする音声認識装置。

【請求項18】 前記句部分文法は、会話状態に基づく句部分文法の集合から選択されることを特徴とする請求項17の装置。

【請求項19】 前記句部分文法は、音声サンプルのコーパスを用いたトレーニングプロセスに基づいて導出されたものであることを特徴とする請求項17の装置。

【請求項20】 前記文仮説検証器は、前記信頼性尺度にも基づいて前記文仮説を生成することを特徴とする請求項17の装置。

【請求項21】 前記抽出済みキー句は、意味タグでラベルされることを特徴とする請求項17の装置。

【請求項22】 前記文仮説検証器は、前記意味タグにも基づいて前記文仮説を生成することを特徴とする請求項21の装置。

【請求項23】 前記文仮説生成器は、前記信頼性尺度にも基づいて前記文仮説を生成することを特徴とする請求項22の装置。

【請求項24】 前記文仮説生成器は、前記信頼性尺度、前記意味タグ、および前記所定の意味情報に基づいて、最も確からしい文仮説を判定することを特徴とする請求項23の装置。

【請求項25】 前記キー句抽出器は複数の部分語モデルに基づいて動作し、前記抽出済みキー句は部分語の列からなることを特徴とする請求項17の装置。

【請求項26】 前記部分語モデルは隠れマルコフモデルからなることを特徴とする請求項25の装置。

【請求項27】 前記キー句検証器は、部分語モデルの集合と、対応する反部分語モデルの集合に基づいて動作することを特徴とする請求項25の装置。

【請求項28】 前記部分語モデルおよび前記反部分語

モデルは隠れマルコフモデルからなることを特徴とする請求項27の装置。

【請求項29】 前記文仮説検証器は、文仮説に対して音響的検証を実行することを特徴とする請求項17の装置。

【請求項30】 前記文仮説検証器は、文仮説に対して意味的検証を実行することを特徴とする請求項17の装置。

【請求項31】 前記文仮説検証器は、最も確からしい1つの文仮説を選択することを特徴とする請求項17の装置。

【請求項32】 前記検証済み文仮説に基づいて意味フレームを生成する意味フレーム生成器をさらに有することを特徴とする請求項17の装置。

【発明の詳細な説明】

【0001】
【発明の属する技術分野】本発明は、音声認識の分野に關し、特に、会話音声を理解する方法に關する。

【0002】

【従来の技術】過去数年間、会話音声の認識および理解のためのシステムが開発され、いくつもの「実世界」の応用で評価されている。いくつかのアプリケーションは、第1のアプリケーションは決定性有限状態文法 (FSG) である。これは、簡単なタスクあるいはアプリケーションに限定されるが、ユーザの発話を受け取る (そしてそれにより認識し最終的に理解する) ものである。このようなシステムでは、認識器は、音声入力全体を、可能な (すなわち、固定した文法により許容される) 単語列のいずれかに合うものを見つける (番号する) ことを試みる。

【0003】
【発明が解決しようとする課題】 実例には、このような固定した文法を使用することは、ほとんど文法的に正しい文 (文法内の文) がシステムに与えられる場合には有効である。しかし、多数のユーザに適用されるほとんど典型的な「実世界」の環境では、さまざまな発話に遭遇し、その多くはこのようなタスクベースの文法によって十分に対応することができない。このような文法的に正しくない文 (文法外の文) には、例えば、無関係な語、口ごもり、反復および予断外の表現などがある。日付あるいは時刻の音声認識のような明らかに単純なサブタスクの場合でさえ、自然なユーザ発話の20%以上が文法外となる可能性がある。システムは、このような条件下で、これらの文法を用いるシステムの性能は低い。そしてこの低い性能は、実行期間中に文法を念入りに調整したにもかかわらず起こる。タスクがさらに複雑な場合にも同様に発生し、これはさらに悪くなる。このような複雑な場合に対処する固定したタスクベースの文法を書いてから十分に調整することは、在外な量の (人間の) 時間および労力を必要とすることが明らか

(3) 特開平10-116094
なっている。
【0004】 上記の問題点は、固定した文レベルの文法を仮定して、その文法が入力全体に適合 (マッチング) しなければならないという一連の要求条件を適用する復号の枠組みから生じている。(発話の文法外部分に適合する) 「フィルタ」モデルの使用は、固定文法のほとんど使う音声サンプルには限定された成功を収めているが、固定文法に固有の基本的な問題点を解決していないために、多くの一般的な発話が認識されないままである。

【0005】 会話音声の認識および理解に対するもう1つのアプローチは、統計言語モデルの使用に關するものである。このようなモデルは、固定した特定の文法に基づくのではなく、大量のサンプルデータを用いた学習 (トレーニング) の結果として統計的に生成された文法に基づく。例えば、米国政府 (ARPA) がスポンサーとなっているATIS (航空旅行情報システム (Air Travel Information System)) プロジェクトは、統計言語モデルのアプローチを用いた会話音声処理に対する包括的プロジェクトである。(例えば、D. A. Dahl, "Expanding the Scope of the ATIS Task: The ATIS-3 Corpus", Proc. ARPA Human Language Technology Workshop, pp. 43-48, 1994, 参照。) この場合、かなりの量のデータが収集され、文法外の発話を認識する能力に關しても、統計言語モデルの使用は比較的成功を収めた。

【0006】 しかし、一般的な「実世界」のアプリケーションでは、データ収集作業自体が大量の (人間の) 時間および労力を必要とするため、統計モデルをトレーニングするためにこのような大量のデータを提供することは実行的でないことが多い。ほとんどのアプリケーションでは、与えられたタスクに対して単純な2連語 (bigram) 言語モデルを構築するのに十分な量のデータを収集するのでさえ実例的ではない。(さらに、注意すべき点であるが、ATISシステムの場合、実行されたシナリオおよびデータ収集はやや人工的であり、従って、実世界の自然発話に固有の問題点を必ずしも反映していない可能性がある。) これらの理由から、「実世界」のアプリケーションに配備され試験されているほとんどの会話システムは、文法外発話を処理する能力が制限されているにもかかわらず、実際には上記のような決定性有限状態文法を使用している。

【0007】 会話音声認識問題のために与えられているもう1つのクラスのアプローチは「単語ボットینگ」方式に關するものである。これは、例えば、R. C. Rose, "Keyword Detection in Conversational Speech Utterances Using Hidden Markov Model Based Continuous Speech Recognition", Computer Speech and Language, 9(9):309-333, 1995, および H. Tsuboi and Y. Takebayashi, "A Real-time Task-oriented Speech Understanding System Using Keyword-spotting", Proc. IE

EE-ICASSP, volume 1, pp. 197-200, 1992, に記載されている。これらのアプローチは、入力発話の非キーワード部分のモデル化に使用する方法に依存して2つのカテゴリに分類される。

【0008】その第1のカテゴリに分類することができざる単語スポットティング方式は、大語彙を認識する能力に基づいているものである。この方式の例としては、J. R. Ro hlback et al., "Phonetic Training and Language Modeling for Word Spotting", Proc. IEEE-ICASSP, volume 2, pp. 459-462, 1993, および、M. Weintraub, "Keyword Spotting Using SRI's DECIPIER Large-Vocabulary Speech-Recognition System", Proc. IEEE-ICASSP, volume 2, pp. 463-466, 1993, に記載されているものがある。この技術は、できるだけ多くの語彙知識を組み込める。キーワードモデルとともに、多くの非キーワード語彙単語モデルをシステムで利用可能とする。しかし、この技術でも、自然音でしばしば見られるロゴモリや自己反復のような非連続的な現象を十分にモデル化することができない。すなわち、すべての一般的なアプローチに固有の問題点を解決していない。さらに、大語彙自然音認識技術は、タスク領域が限定される場合には特に、性能および効率性に問題がある。

【0009】単語スポットティング方式の第2のカテゴリは、入力発話の非キーワード部分をモデル化するため用いられる部分語(subword)モデルの並列ネットワークまたは単語ガジェットとともに単純(すなわち限定語)単語スポットティングを使用するものである。このような単語スポットティングは、J. G. Wilpon et al., "Automatic Recognition of Keywords in Unconstrained Speech Using Hidden Markov Models", IEEE Trans. Acoust., Speech & Signal Process., 38 (11): 1870-1878, 1990, および、R. C. Rose and D. B. Paul, "A Hidden Markov Model Based Keyword Recognition System", Proc. IEEE-ICASSP, pp. 129-132, 1990, に記載されているものがある。残念ながら、ガジェットモデルも、部分語モデルの並列ネットワークも、非キーワードに含むものを見つけて性能が十分ではなく、そのため、キーワードモデルが発話の無関係(すなわち非キーワード)部分と誤って合わせられることも多い。この結果、多くの「フォールスアラーム」(すなわち、キーワードの誤った「認識」)が起こる。さらに、このカテゴリに属する既存のほとんどのシステムは、語彙に依存してキーワードモデルおよびガジェットモデルを「調整」し、それにより、部分語に基づいた単語認識アプローチの利点の多くを犠牲にしている。これらの理由から、このカテゴリの単語スポットティング方式で応用が成功しているのは、例えば数字音声認識のタスクのような非常に小さい語彙を含むタスクのみである。

【0010】

【課題を解決するための手段】本発明の発明者が認識し

たところでは、ほとんどの会話音声発話(すなわち「文」)はタスクに閱するあるキーワードおよび「キーワード(キーフレーズ)」を含む。その認識により発話の部分的あるいは全体的な理解が可能となる一方で、発話の他の部分は文脈にはタスクに関係がないので無視すべきである。(注意すべき点であるが、「文」という用語は、本明細書では、任意の単語列を意味し、そのような単語列が文法的に正しい文構造を有するかどうかは問わない。また、「キーワード」という用語の使用は、本明細書では、1個以上の単語からなる列を含むものとする。すなわち、キーワードは単一の単語からなる「キーワード」である。)すなわち、自由度の高い音声理解システムは、文の文脈的に重要な部分を検出し無関係な部分を拒絶するアプローチに基づいて構成することができ、従来の文法的制約を緩和し、かつ、認識されるキーワードの集合に特に注目することによって、例えば固定した形式的な文法に基づいて受容されるよりも多くの発話が受容される。

【0011】そこで、本発明の実施例によれば、自由度の高い(すなわち、制約のない)音声の理解を実現するために使用可能なキーワードの検出および発話の技術が実現される。具体的には、単語列(すなわち文)からなる音声発話に「多重パス」手続きが適用される。まず、例のキーワードを検出(すなわち、認識)する。次に、これらのキーワードに信頼性尺度を割り当て、その信頼性尺度を正しい値と比較することによってそれらのキーワードを証明し、その結果として、検証済みキーワード候補の集合を得る。次に、検証済みキーワード候補を、所定の(例えばタスク固有の)意味情報に基づいて結合して文脈を検証する。文脈は、個々のキーワードの信頼性尺度に基づいて生成することも可能である。最後に、これらの文脈を検証して、検証済み文脈を検出し、その結果、音声発話の理解を得る。

【0012】さらに、会話ベースのシステムでは特に、大規模なタスク内の会話の特定の状態に応じて(すなわち、「サブタスク」に対して)、個別の句部分文法を使用することが可能である。例えば、会話ベースの自動音声タスク内では、システムは、与えられた時点において、要求された単語が必要となる日時を検出する必要があり、この場合、予測される応答は、時間的な情報のみを与えるものであると限定することができ、自由度の高い会話ベースのシステムは、文脈を少なくとも部分的に理解することができ、さらに、会話セクションが進むうちに、必要な明確化(曖昧さの除去)を実行することも可能である。

【0013】

【発明の実施の形態】

【はじめに】本発明の実施例によれば、会話音声の認識

および理解のためのシステムは、(例えば、非キーワード大語彙知識を用いることなく)部分語ベースの音声認識の一般的な枠組みで、無関係部分を除外して「認識」せずに、発話の重要部分を認識することによって実現される。(部分語ベースの音声認識は、当業者には周知であるが、音節、半音節あるいは音素のような単語セグメントのモデリングおよびマッピングを含む。次に、それらの単語セグメント(すなわち、部分語)の列に、語彙内の各単語をマッピングするために、辞書(lexicon)が提供される。こうして、単語に対応するモデルは、実質的に、辞書によって指定される。その単語を構成する部分語のモデルの連接からなる。)図1に、本発明の実施例による音声認識および音声発話の理解を実行する1つの例示的なシステムの図を示す。

【0014】注意すべき点であるが、従来技術の最も重大な問題点のうちの1つは、従来の音声認識装置は一般に、その結果にどのくらいの信頼性をおくことができるかが分からないことである。この理由で、図1に示した本発明の実施例によれば、認識した結果に対する信頼性を定実行し、それに信頼性尺度を割り当てて検証方法を用いる。(例えば、R. A. Sukkar et al., "A Vocabulary Independent Discriminatively Trained Method for Rejection of Non-Keywords in Subword-Based Speech Recognition", Proc. EuroSpeech-95, pp. 1629-1632, 1995, R. A. Sukkar et al., "Utterance Verification of Keyword Strings Using Word-Based Minimum Verification Error (WBE-MVE) Training", Proc. IEEE-ICASSP, pp. 518-521, 1996, および、M. Rahim et al., "Discriminative Utterance Verification Using Minimum String Verification Error (MSVE) Training", Proc. IEEE-ICASSP, 1996, 参照。)このような発話検証法を図1の実施例のシステムに統合することによって、キーワード(あるいは、この場合にはキーワード)の検出の信頼性を高くすることができる。すなわち、キーワードモデルへの正しくないマッピングがすなわち「フォールスアラーム」は大幅に減少する。

【0015】また、図1の実施例のシステムは、このように「フォールスアラーム」をさらに減少させる。システムは、このようなキーワード(あるいはキーワード)マッピングおよび検証プロセスの世帯の結果として「最終判定」をしない。むしろ、検証したキーワードあるいはキーワードの組み合わせ(すなわち、文)に基づいて意味解析を実行して文脈を検出し、それを別の検証プロセスで検証する。特に、この文脈検証プロセスは、会話内にあるいくつかの部分語からなる「部分入力」で実行される。

【0016】既に指摘したように、図1の実施例のシステムは、検出単位として、キーワードのみを使用するのではなく、キーワードを使用する。上記の単語スポットティング方式は一般に、局所的ノイズや乱雑な音によって容易

にトリガされる小さいテンプレートを使用する。より長い検出単位(すなわち、単なるキーワードの代わりにキーワード)を使用することは、より特徴的な情報を含むことになり、その結果、認識精度および検証精度の両方、より安定な音響マッチングが得られるので、有効である。

【0017】具体的には、キーワードは、1個以上のキーワードと、おそらくは、機能語との列からなる。例えば、「In the morning」は、期間についてのキーワードであり、「in downtown Chicago」は、地理的場所についてのキーワードである。このような句は、自然音で発話されるときでも、一般に思惟なしで発話される。

【0018】ここに記載する本発明の実施例によれば、検出されたキーワードには概念情報タグが付けられる。実際には、キーワードは、例えば時刻および場所のような、意味(セマンティック)フレームにおける意味スロットに直接対応するように定義される。(意味フレームは、当業者には周知の用語であるが、与えられたアプリケーションに対して、全語によって部分的にあるいは完全に充満される情報テンプレートからなる。)従来のn語長(n-gram)言語モデルによって定義されるようなボトムアップ句(例えば、B. Suhm and A. Waibel, "Towards Better Language Models for Spontaneous Speech", Proc. IC SLP, pp. 831-834, 1994, E. P. Giachin, "Phrase Bigrams for Continuous Speech Recognition", Proc. IEEE-ICASSP, pp. 225-228, 1995, および、S. Deigne and F. Bimbot, "Language Modeling and Evaluation of Multigrams", Proc. IEEE-ICASSP, pp. 169-172, 1995, 参照。)とは異なり、本実施例によって認識されるトップダウンキーワードは、容易に意味表現へと直接にマッピングされる。従って、これらのキーワードの検出は、直接に、発話の確実な理解につながる。

【0019】具体的には、図1の実施例のシステムは、キーワード検出器11、キーワード検出器12、文脈検出器13および文脈検出器14を有する。特に、キーワード検出器11は、会話状態に特有の句部分文法(すなわち、キーワード文法21)の集合を用いてキーワードの集合を認識するための部分語ベースの音声認識器からなる。検出されたキーワードは、次に、意味(セマンティック)タグが付けられる。このタグは、文脈検出器13(後述)によってその後に行われる文レベルの解析で有用となる。キーワード検出器11によって用いられる部分語モデル認識器は、辞書23および部分語モデル22を使用する。これは、例えば、当業者には周知の従来の最小分類誤差(MCE (minimum classification error))基準に基づいてトレーニングされたものである。これらのモデル自体は、例えば、同じく当業者には周知の隠れマルコフモデル(HMM)からなることも可能である。

【0020】次に、検出されたキーワードは、キーワード検出器

12によって検証され、信頼性尺度が割り当てられる。上記のように、このプロセスは、これがなければ起こり得る多くのフォールスアラームを除去する。実施例のキリ一句検証器12は、当業者に周知の「反部分語モデル」を用いて、距離されたキリ一句の各部分語をテストする。部分語レベルの検証の組合せからなる。キリ一句検証器12は辞書23、部分語モデル22および反部分語モデル24を使用する。これらは、例えば、最小検証誤差(MVE (minimum verification error))基準を用いてトレーニングされたものである。

【0021】図1の実施例の第3の構成要素は文脈生成器13である。これは、例えばタスク固有の意味情報25を用いて、検証されたキリ一句候補を、1つの文脈説へと結合する。例えば、I. Kawahara et al., "Concept-Based Phrase Spotting Approach for Spontaneous Speech Understanding", Proc. IEEE-ICASSP, pp. 291-294, 1996, に記載されたようなスタック複合器を用いて、意味制約を満たす最適な仮説を探索することができ

る。

【0022】最後に、文脈候補検証器14によって、音響的かつ意味的に最も意味仮説が検証される。最終出力(すなわち、少なくとも1つの検証された文脈)が生成される。文脈候補検証器14は、意味情報25、辞書23、部分語モデル26および反部分語モデル27を使用する。キリ一句に付けられた意味タグが、キリ一句検証器1によって提供された意味仮説生成器13によって使用されるため、検証された文脈は本質的に、直接に対応する「意味」を有し、それにより、個々のアプリケーションに必要に応じて意味フレーズの生成が可能となる。

【0023】キリ一句抽出11は、キリ一句抽出を実行する。これは、会話状態に依存する特定のサブタスクに基づくことが可能である。具体的には、各サブタスクごとに、キリ一句パターンが1つ以上の決定性有限状態文法として記述される。これは、実施例では、キリ一句抽出器11によってキリ一句文法21から選択される。これらの文法は、タスク仕様から直接に人手により導出することも可能であり、あるいは、当業者に周知の従来の学習手続きを用いて、小さいコーパスから自動的にまたは半自動的に(すなわち、人間の支援のもとで)生成することも可能である。

【0024】一般に、キリ一句は、従来のキリ一句に比べて、"at the"や"near"のような機能語を含む。これにより、従来のキリ一句のみのマッチングに比べて、より安定なマッチングが可能となり、抽出精度が改善される。(例えば、前掲J. Kawahara et al., "Concept-Based Phrase Spotting Approach for Spontaneous Speech Understanding"を参照。) いずれのキリ一句にも含まれないが、キリ一句に伴う冗項句も検証され、埋め込まれたキリ一句を含むパターンを形成するために使用

される。

【0025】特に、キリ一句および冗項句の文法はネットワークへとコンパイルされる。このネットワークにおいて、キリ一句は繰り返して現れ、ガーベジモデルがキリ一句の出現の間に埋め込まれる。しかし、注意すべき点であるが、単純な繰り返して埋め込まれる可能性がある。例えば、目の繰り返して埋め込まれる場合、"twenty four"と"twenty"を区別することはできない。従って、不可能なキリ一句の結合を禁止する追加の制約も組み込む必要がある。

【0026】従って、抽出ユニットは、許容される結合および仮説を有するキリ一句部分文法オートマトンのネットワークからなる。このようなオートマトンは、結合重みを評価することによって、確率的言語モデルへと容易に拡張することができ、このようなモデルを使用することにより、文レベルの文法と比べてあまり複雑にならない、適用範囲が広がる。

【0027】例として、図2に、単純化した(すなわち、簡略化した)句ネットワークの例を示す。これは、「データ取得」サブタスクに適用された場合に、図1の実施例のシステムのキリ一句抽出器11によって使用されることが可能である。このネットワーク例の完全な実現により、曜日、月、日、および年の要素は11の要素の反復が、適当な制約のもとで許容される。(このような完全な実現の全長は99語である。) この特定のサブタスクでは、キリ一句は組み込まれない。

【0028】さらに具体的には、ここに記載する本発明の実施例によって採用されている抽出方法は、フローワードバックワード2パス探索に基づくものである。これは、例えば、W. Chou et al., "An Algorithm of High Resolution and Efficient Multiple String Hypothesis Resolution for Continuous Speech Recognition Using Inter-Word Models", Proc. IEEE-ICASSP, volume 2, pp. 53-56, 1994, に記載されている。本発明の別の実施例では、代わりに、当業者に周知の1パス抽出法を使用することも可能である。

【0029】A. 読者スタック複合器 (例えば、前掲のI. Kawahara et al., "Concept-Based Phrase Spotting Approach for Spontaneous Speech Understanding"に記載されているもの) は、N番目までの最良候補として得られるN個の最良仮説は一般に、1〜2語が重なり、残基全体に基づいてストリング仮説を生成することではなく) スパ域結の1語に基づいてキリ一句候補を生成することであるので、仮説を延長しても既に延長された仮説と同じ仮説になる場合にはその仮説は捨てられる。

【0030】特に、本実施例のスタック複合器は、キリ一句ネットワークのマーキング(emerging)状態にマークを付けることによって実現される。当業者には周知のよう

に、マーキング状態は、キリ一句あるいは冗項句が終了し、さらに延長する次の(すなわち新たな)句の最初に挿入することになるノードに対応する。

【0031】スタック複合器に对应する「ポインタ」される仮説に、出力されるべき完全な句であるというタグが付いている場合、本発明の手続きは、もう1語だけその句を延長し、その句を最良候補と並べる。このノードに、以前のいずれかの仮説が同じ時点に到達している場合、抽出した句を出力した後には現在の仮説は捨てられる。そうでなければ、その時点は、その後の探索のためにマークされる。

【0032】注意すべき点であるが、この抽出手続きは、冗長な仮説延長のない効率的なものであり、スコア順に、正しいN番目までの最良のキリ一句候補を生成する。本発明のさまざまな実施例によれば、この手続きは、所望の個数の句を生成したことに基づいて、あるいは、あるスコアしきい値に基づいて、終了することも可能である。例えば、仮説のスコアが、最良スコア仮説の0.99倍より小さい値に到達したときに、抽出を終了することも可能である。

【0033】「キリ一句検証と信頼性尺度」図1の実施例のシステムのキリ一句検証器12は、部分語レベルのテストに基づいて、抽出された句の検証を行う。具体的に、与えられた句の各部分語nに対して、検証スコアは、次のような従来の尤度比(LR (likelihood ratio)) テストに基づいて計算される。

$$\log L R_n = (\log P(O|\lambda_n^*) - \log P(O|\lambda_n)) / 10 \quad (2)$$

注意すべき点であるが、式(2)の第1項は認識スコアそのものである。上記の計算の結果は単に、計算されるスコアを反部分語モデルのスコアで除き、その結果を正規化することである。

【0036】キリ一句検証器12は、抽出された各キリ一句ごとに、対応する部分語レベルの検証スコアを組み合わせる。

$$CM = \left(\log L R_1, \dots, \log L R_n \right) \quad (3)$$

信頼性尺度(CM)が、ある所定のしきい値を超える場合に、与えられたキリ一句は承認される。実施例では、しきい値の値は、例えば-0.15に設定される。

【0037】本発明のさまざまな実施例において、さまざまな信頼性尺度関数を使用することができ、例えば、第1の明示的な信頼性尺度CMは、フレーズ継続★

$$CM_1 = \frac{1}{L} \sum_{i=1}^L (a_i \cdot \log L R_n) \quad (4)$$

上記の式で、 L は、部分語nの継続時間を表し、 L は句の全継続時間である。すなわち、 $L = \sum L_n$ である。【0038】第2の明示的な信頼性尺度CMは、部分語セグメントによる正規化に基づく。特に、これは、与えられたキリ一句のすべての部分語の対数尺度比の単なる平均である。(一実施例では、句セグメンテーション後

* $L R_n = P(O|\lambda_n^*) / P(O|\lambda_n)$ (1)
ただし、Oは、観測フレーズの列を表し、 λ_n^* および λ_n は、それぞれ、部分語nに対する正しい部分語モデルおよび反部分語モデルを表す。(部分語モデルは部分語モデル22から得られ、対応する反部分語モデルは反部分語モデル24から得られる。) 認識の結果として、観測Oは、部分語nに対して、ビタビアルゴリズムにより与えられる。観測P (O | λ_n^*) および P (O | λ_n) が得られる。(ビタビアルゴリズムは、当業者に周知の従来のスコアリング方法である。)

【0034】各部分語モデルに対して、対応する反部分語モデルは、逆同しまたは部分語クラスをまとめること(クラスタ化)によって構成される。各反部分語モデルは、対応する部分語モデルと同じ構造、すなわち、同じ個数の状態およびミクスチャを有する。反部分語モデルは、特定の部分語の検証専用であるため、反部分語モデルをリファレンスとして使用して復号を行うことにより、部分語モデルの無制約な復号を行うのに比べて、弁別性が改善される。こうして、システムは、認識器によってなされる復号誤りを拒絶する能力が増大する。この(検証)ステップでは、文脈独立な反部分語モデルを使用することも可能であるが、認識ステップは、文脈依存の部分語モデルを用いて実行される。

【0035】特に、上記の式(1)の対数をとり、その結果を、観測Oの継続時間長 L に基づいて正規化することにより、量 $\log L R_n$ が次のように定義される。

$$\log L R_n = (\log P(O|\lambda_n^*) - \log P(O|\lambda_n)) / 10 \quad (2)$$

※ 量 $\log L R_n$ は、信頼性尺度(CM (confidence measure))を計算する。例えば、抽出されたキリ一句がN個の部分語を含む場合、このキリ一句に対する信頼性尺度は、対応するN個の尤度比の関数とすることが可能である。具体的には、次のようになる。

$$CM = \left(\log L R_1, \dots, \log L R_n \right) \quad (3)$$

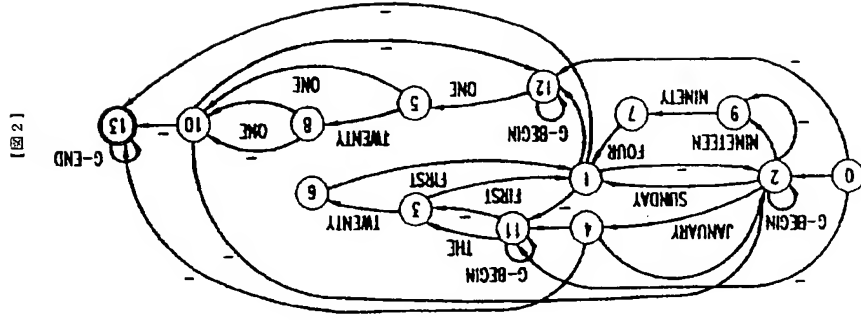
★ 時間による正規化に基づく。特に、これは、正しい部分語モデルに対して得られるビタビスコアと、対応する反部分語モデルに対して得られるビタビスコアの差に等しい。すなわち、次のようになる。

$$CM = \left(\log L R_1, \dots, \log L R_n \right) \quad (4)$$

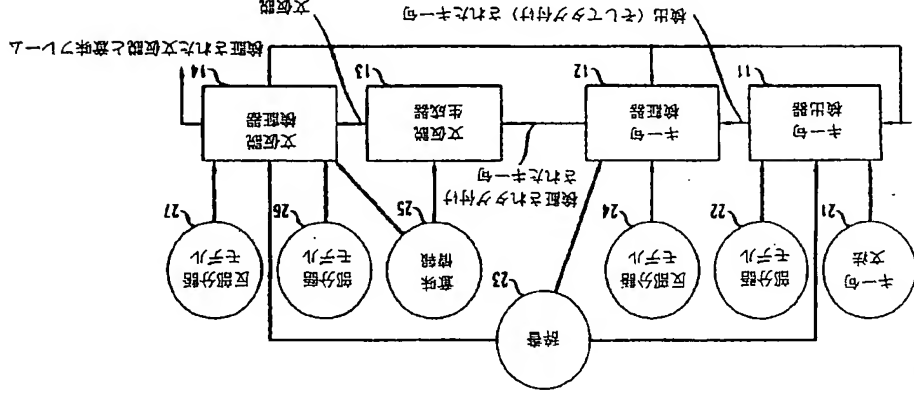
【数1】

に単語間文脈情報が行われるため、最後の部分語に対して特別の考慮がなされる。) すなわち、次のようになる。

【数2】



【図1】



【図2】「日付取得」サブタスクに適用した場合に、図1の例示的なシステムによって使用されることが可能な単純化された句ネットワーク例の図である。

【符号の説明】

- 11 キー句抽出器
- 12 キー句検出器
- 13 文反訳生成器
- 14 文反訳検出器
- 21 キー句文法
- 22 部分語モデル
- 23 辞書
- 24 反部分語モデル
- 25 意味情報
- 26 部分語モデル
- 27 反部分語モデル

(RAM) からなることが可能である。超大規模集積回路(VLSI)ハードウェアや、カスタムVLSI回路を適用プロセッサやDSP回路と組み合わせたものも可能である。

【0054】また、「キー句抽出器」、「キー句検出器」、「文反訳生成器」、および「文反訳検出器」という用語は、対応する機能を実行する任意のメカニズムを含む。

【0055】以上述べたごとく、本発明によれば、図1に示した形式的な文法に基づいて受容されるよりも多くの発話を受容される、効率的および自由度の高い会話音声認識が実現される。

【図面の簡単な説明】
【図1】本発明の実施例による音声認識および音声発話の理解を実行するシステムの図である。

フロントページの続き

(71)出願人 596077259
600 Mountain Avenue,
Murray Hill, New Jersey 07974-0636 U. S. A.

(72)発明者 ビン・ホウ・ジョン
アメリカ合衆国、07059 ニュージャージー
一、ウォレン、サクス レーン 8

(72) 発明者	かわはら たつや	(12)	特開平 10-116094
	京都府京都市伏見区東奉行伏見御堂122		
(72) 発明者	チン-フイ リー		
	アメリカ合衆国、07574 ニュージャージー		
	ー、ニュー プロビデンス、ラニーメデ		
	パークウェイ 118		

NETWORK INTERACTIVE USER INTERFACE USING SPEECH RECOGNITION AND NATURAL LANGUAGE PROCESSING

Publication number: JP2002524806T

Publication date: 2002-08-06

Inventor:

Applicant:

Classification:

- International: G06F3/16; G06F13/00; G06F17/27; G06F17/28; G10L15/18; G10L15/26; H04L29/06; G06F3/16; G06F13/00; G06F17/27; G06F17/28; G10L15/00; H04L29/06; (IPC1-7): G06F17/28; G06F3/16; G06F13/00

- European: G06F3/16; G06F17/27; G10L15/18C2; G10L15/26A; H04L29/06

Application number: JP20000569391T 19990908

Priority number(s): US19980150459 19980909; US19980166198 19981005; WO1999US20447 19990908

Also published as:

WO0014728 (A1)
EP1110206 (A1)
US6434524 (B1)
EP1110206 (A0)
CA2343150 (A1)

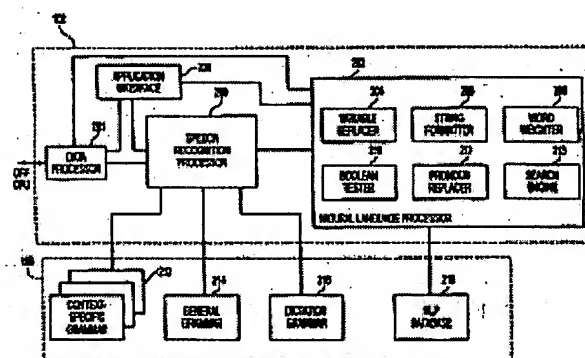
more >>

Report a data error here

Abstract not available for JP2002524806T

Abstract of corresponding document: WO0014728

A system and method for interacting with networked objects, via a computer using utterances, speech processing and natural language processing. A Data Definition File relates networked objects and a speech processor. The Data Definition File encompasses a memory structure relating the networked objects, including grammar files and a natural language processor. The speech processor searches a first grammar file for a matching phrase for the utterance, and for searching a second grammar file for the matching phrase if the matching phrase is not found in the first grammar file. The system also includes a natural language processor for searching a data base for a matching entry for the matching phrase; and an application interface for performing an action associated with the matching entry if the matching entry is found in the database. The system utilizes context-specific grammars, thereby enhancing speech recognition and natural language processing efficiency. Additionally, the system adaptively and interactively "learns" words and phrases, and their associated meanings.



Data supplied from the esp@cenet database - Worldwide

【特許請求の範囲】

【請求項 1】 ネットワーク・オブジェクトと関連付けられたネットワーク・オブジェクト・テーブルをネットワークを介して転送する工程と、上記ネットワーク・オブジェクトに対する一致エントリに関して上記ネットワーク・オブジェクト・テーブルを検索する工程と、上記ネットワーク・オブジェクト・テーブル内において上記一致エントリが発見されたなら上記一致エントリに関連付けられた動作を実施する工程と、を含む、ネットワーク・オブジェクトとの音声対話の為にコンピュータを更新する方法。

【請求項 2】 前記ネットワーク・オブジェクト・テーブルはウェブ・サイトから読み出される、請求項 1 記載の方法。

【請求項 3】 前記ネットワーク・オブジェクト・テーブルは、複数のネットワーク・オブジェクトに対して各ネットワーク・オブジェクト・テーブルを記憶する箇所から読み出される、請求項 1 記載の方法。

【請求項 4】 前記ネットワーク・オブジェクト・テーブルは、文脈固有文法も含むダイアログ定義ファイル内に含まれる、請求項 1 記載の方法。

【請求項 5】 前記ネットワーク・オブジェクト・テーブルは、自然言語プロセッサ・データベースに対する各エントリも含むダイアログ定義ファイル内に含まれる、請求項 1 記載の方法。

【請求項 6】 前記ネットワーク・オブジェクト・テーブルは、文脈固有文法および自然言語プロセッサ・データベースに対する各エントリも含むダイアログ定義ファイル内に含まれる、請求項 1 記載の方法。

【請求項 7】 ネットワーク・オブジェクトと関連付けられたダイアログ定義ファイルであってネットワーク・オブジェクト・テーブルを含むダイアログ定義ファイルを転送するネットワーク・インタフェースと、

上記ネットワーク・オブジェクトに対する一致エントリに関して上記ネットワーク・オブジェクト・テーブルを検索するデータ・プロセッサと、

上記ネットワーク・オブジェクト・テーブル内において上記一致エントリが発見されたなら上記一致エントリに関連付けられた動作を実施するアプリケーション

Equivalent to Prior Art Literature

(11) 特許出願公表番号

特表 2002-524806

(P2002-524806A)

(43) 公表日 平成 14 年 8 月 6 日 (2002.8.6)

(12) 公表特許公報 (A)

(19) 日本国特許庁 (JP)

(51) Int. Cl.	識別記号	FI	チコード (参考)
G 0 6 F	17/28	G 0 6 F	17/28
	3/16		3/16
	13/00		13/00
			3 4 0 A
			3 4 0 N
			5 5 0 A

審査請求 未請求 予備審査請求 有 (全 49 頁)

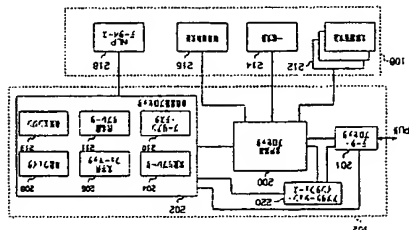
(21) 出願番号	特願 2000-569391 (P2000-569391)	(71) 出願人	ワン ボイス テクノロジーズ インコーポレーテッド
(86) (22) 出願日	平成 11 年 9 月 8 日 (1999.9.8)		アメリカ合衆国カリフォルニア州 92122
(85) 優先権主張日	平成 13 年 3 月 9 日 (2001.3.9)		サン ディエゴ スイート 240 グリーン
(86) 国際出願番号	PCT/US99/20447		ウィッチ ドライブ 6333
(87) 国際公開番号	WO00/14728 Prior art literature		ウィーバー、ディーン シー
(87) 国際公開日	平成 12 年 3 月 16 日 (2000.3.16)	(72) 発明者	アメリカ合衆国 92131 カリフォルニア州 サンディエゴ サイバラス テラス
(31) 優先権主張番号	09/150,459		ブレイス 11422
(32) 優先日	平成 10 年 9 月 9 日 (1998.9.9)	(74) 代理人	井理士 豊田 博宣 (外 1 名)
(33) 優先権主張国	米国 (US)		ファーム (参考) 55091 MA15 CA12 CA14 CA32 CA03
(31) 優先権主張番号	09/166,198		DM3
(32) 優先日	平成 10 年 10 月 5 日 (1998.10.5)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 音声認識および自然言語処理を使用したネットワーク用対話ユーザ・インタフェース

(57) 【要約】

発願、音声処理および自然言語処理を使用してコンピュータを介してネットワーク・オブジェクトと対話するシステムおよび方法。データ定義ファイルは、ネットワーク・オブジェクトおよび音声プロセッサを関連付ける。データ定義ファイルは、文法ファイルを含むネットワーク・オブジェクトを自然言語プロセッサに対して関連付けるメモリ構造を包含する。音声プロセッサは、上記発願に対する照会対象語句に関して第 1 文法ファイルを検索し、もし上記第 1 文法ファイル内で上記照会対象語句が見えられれば上記照会対象語句に関して第 2 文法ファイルを検索する。上記システムはまた、上記照会対象語句に対する一致エントリに関してデータベースを検索する自然言語プロセッサ; および、もし上記一致エントリが上記データベース内で発見されたら一致エントリに関連付けられた動作を実施するアプリケーション・インタフェースも含む。上記システムは文脈固有文法を利用することにより、音声認識および自然言語処理の効率を高める。更に上記システムは、単語および語句ならびにそれらの関連意味を適時的かつ対話的に “学



ン・インタフェースと、

を含む、ネットワーク・オブジェクトとの音声対話の為にコンピュータを更新するシステム。

【請求項8】 ネットワーク・オブジェクトと関連付けられたダイアログ定義ファイルであってネットワーク・オブジェクト・テーブルを含むダイアログ定義ファイルを発見する工程と、

上記ダイアログ定義ファイルを読み出す工程と、

上記ネットワーク・オブジェクトに対する一致エントリに関して上記ネットワ

ーク・オブジェクト・テーブルを検索する工程と、

上記ネットワーク・オブジェクト・テーブル内において上記一致エントリが発見されたなら上記一致エントリに関連付けられた動作を実施する工程と、

を含む、ネットワーク・オブジェクトとの音声対話の為にコンピュータを更新する方法。

【請求項9】 前記ダイアログ定義ファイルはウェブ・サイトから読み出される、請求項8記載の方法。

【請求項10】 前記ダイアログ定義ファイルは複数のネットワーク・オブジェクトに対するダイアログ定義ファイルを記憶する箇所から読み出される、請求項8記載の方法。

【請求項11】 前記ダイアログ定義ファイルは記憶媒体から読み出される、請求項8記載の方法。

【請求項12】 前記ダイアログ定義ファイルは文脈固有文法を含む、請求項8記載の方法。

【請求項13】 前記ダイアログ定義ファイルは自然言語プロセッサ・データベースに対するエントリを含む、請求項8記載の方法。

【請求項14】 前記ダイアログ定義ファイルは文脈固有文法と自然言語プロセッサ・データベースに対するエントリとを含む、請求項8記載の方法。

【発明の詳細な説明】

【0001】

(発明の背景)

1. 発明の分野

本発明はネットワーク・コンピュータ用ユーザ・インタフェースの音声認識に関する。より詳細には本発明は、音声認識および自然言語処理を使用してユーザとコンピュータが対話する新規な方法およびシステムに関する。本発明は、1998年9月10日出願されて“音声認識および自然言語処理を使用した対話型ユーザ・インタフェース”と称された米国特許出願第09/150,459号の一部継続出願である。

【0002】

1.1. 関連技術の説明

コンピュータが普及するにつれ、コンピュータの理解およびコンピュータとの対話が多くの人々に取り困難であることが明らかとなっている。コンピュータ操作の為にユーザは、原始的なコマンドおよび直感的でない手順を学ぶべき場合が多い。例えば殆どのパーソナルコンピュータは、概してメニュー主導であるウィンドウズ（登録商標）式のオペレーティング・システムを使用する。故にユーザは、どのメニュー・コマンドもしくは一連のコマンドが所望結果を与えるのかを学ぶ必要がある。

【0003】

更に、コンピュータに対する習用の対話はキーボードもしくはマウスなどの手動入力デバイスの故に遅くなることも多い。多くのコンピュータ・ユーザは熟練タイピストではない。その結果、これらの手動入力デバイスによりコンピュータにコマンドおよび単語を入力するには相当の時間が掛かる。故に、コンピュータ、ならびに、ウェブ・サイトなどのネットワーク・オブジェクトと対話するの容易かつ高速で更に直感的な方法が必要なのは明らかである。

【0004】

コンピュータとの対話の為に提案されたひとつの方法は、音声認識である。音声認識は、人間の音声を音響的に検出すると共に検出された音声を一連の単語へ

と翻訳すべく協働するソフトウェアおよびハードウェアを必要とする。当業界において公知である如く音声認識は、ハードウェアが検出した音を、音楽 (phone) と称される分割不能な更に小さな音へと分解することで機能する。各音楽は別個の音の単位である。例えば “those” という単語は3個の音楽から成り、第1の音楽は “th” 音であり、第2の音楽は “o” 音であり、第3の音楽は “s” 音である。音声認識ソフトウェアは、検出された各音楽を記憶辞書からの既知の各単語と一致させるべく試行する。音声認識システムの一例は、1998年11月8日に発行されて “音声認識装置および方法” と称され且つ Dragon Systems, Inc. に譲渡された米国特許第4,783,803号に与えられるが、該米国特許第4,783,803号により本明細書中に援用する。現在では、Dragon Systems, Inc. および International Business Machines Inc. などの会社から市販された多くの音声認識ソフトウェア・パッケージが在る。

【0005】

これらの音声認識ソフトウェア・パッケージもしくはシステムのひとつの制限は、それらが通常はコマンドおよび制御もしくは聴音書取 (dictation) 機能のみを実施することである。故にユーザは依然として、コンピュータを操作する為にはコマンドの用語集を学ぶ必要がある。

【0006】

これらの音声認識システムに対して提案された改良は、検出された単語を自然言語処理システムを用いて処理することである。自然言語処理は一般的に、検出された単語の文法的関係および相対的文脈を分析することにより各単語の概念的 “意味” (例えば、話し手が意図している意味) を決定する必要がある。例えば1989年12月12日に発行されると共に “自然言語テキストに対するパーサ (PARSER)” と称され且つ International Business Machines Inc. に譲渡された米国特許第4,887,212号は言及したことにより本明細書中に援用するが、該米国特許第4,887,212号は言及したことの他に、辞書検索および文法分析を使用して入力単語ストリームを解析する方法を教示している。

【0007】

音声認識とともに自然言語処理を使用すると、話された単語を使用して、キーボードもしくはマウスなどの手動入力よりも強力なコンピュータの操作ツールが提供される。しかし乍ら従来の自然言語処理システムのひとつの欠点は、音声認識システムにより検出された単語の正しい “意味” を自然言語処理システムが決定し得ない可能性があることである。斯かる場合にユーザは通常、次の試行ではその自然言語処理システムが正しい “意味” を決定することを期待してその語句を言い換えたり言い直したりする必要がある。これは、ユーザが文もしくはコマンドの全体を言い直す必要があるときに相当の遅延に繋がるのは明らかである。従来システムの別の欠点は、音声認識に必要な処理時間が法外に長くなり得ることである。これは主として、処理されるべき大量の情報と比較して処理リソースの速度が有限だからである。例えば従来の多くの音声認識プログラムにおいては検索される辞書ファイルのサイズの故に、発語 (utterance) の認識に要する時間は長い。

【0008】

従来の音声認識/自然言語処理システムの更なる欠点は、それらが対話型で無いため新たな状況に対処し得ないことである。未知のもしくは新たなネットワーク・オブジェクトにコンピュータ・システムが遭遇したとき、コンピュータとオブジェクトとの間には新たなリレーションシップが形成される。故に従来の音声認識/自然言語処理システムは、それまで未知であったネットワーク・オブジェクトにより課された新たなリレーションシップに起因する状況には対処し得ない。結果として、コンピュータとの会話形式対話は不可能である。ユーザは、完全な概念をコンピュータに伝達する必要がある。ユーザはまた断片的な文を話すことが出来ない、と言うのも、(以前の発語の意味に依存する) それらの断片的な文の意味が失われるからである。

【0009】

必要なのは、上述の各欠点を回避する音声認識および自然言語処理を活用した、コンピュータとの対話型ユーザ・インタフェースである。

(発明の概要)

本発明は、発語、音声処理および自然言語処理を使用してコンピュータと対話する新規で優れたシステムおよび方法である。概略的に上記システムは音声プロセッサを含むが、該音声プロセッサは、上記発語に対する照合対象語句に関して第1文法ファイルを検索し、且つ、もし上記第1文法ファイル内で上記照合対象語句が発見されなければ上記照合対象語句に関して第2文法ファイルを検索する。上記システムはまた、上記照合対象語句に対する一致エントリに関してデータベースを検索する自然言語プロセッサ、および、もし上記一致エントリが上記データベース内で発見されたならば該一致エントリに関連付けられた動作を実施するアプリケーション・インタフェースも含む。

【0010】

好適実施例において上記自然言語プロセッサは、上記一致エントリが上記データベース内で発見されなければ上記データベース、上記第1文法ファイルおよび上記第2文法ファイルの少なくともひとつを上記照合対象語句により更新する。

【0011】

上記第1文法ファイルは文脈固有文法ファイルである。文脈固有文法ファイルとは、特定主題に相当に関連する単語および語句を含むファイルである。上記第2文法ファイルは一般文法ファイルである。一般文法ファイルとは、文脈に鑑みて解釈の不要な単語および語句を含むファイルである。すなわち、上記一般文法ファイル内の単語および語句は一切の起源文脈に属さない。本発明では上記一般文法ファイルを検索する前に上記文脈固有文法ファイルを検索することから、ユーザは更に会話的な形式でコンピュータと意志疎通でき、その場合に話された単語は、もし上記文脈固有文法ファイル内で発見されれば、最も最近に話された主題に鑑みて解釈される。

【0012】

本発明の更なる側面において上記音声プロセッサは、もし上記照合対象語句が上記一般文法ファイル内で発見されなければ上記照合対象語句に関して聴音書取文法を検索する。上記聴音書取文法は、一般的な単語および語句の大規模な用語集である。上記文脈固有文法および一般文法を最初に検索することにより音声認識時間が相当に短縮されることが期待されるが、これは、上記文脈固有文法およ

び一般文法は上記聴音書取文法よりも物理的に小さいからである。

【0013】

本発明の別の側面において上記自然言語プロセッサは、上記データベースを検索する前に上記照合対象語句の少なくとも一個の単語を置換する。この置換は、上記照合対象語句の少なくとも一個の単語をワイルドカードで置換する上記自然言語プロセッサ内の変数リプレースにより達成され得る。上記語句内の(“単語変数”と称される)一定の単語をワイルドカードで置換することにより、上記データベース内のエントリ数は相対的に減少され得る。更に上記自然言語プロセッサ内の代名詞サブステディュータは上記照合対象語句内の代名詞を固有名称に置換することにより、上記データベース内にユーザ特有事実が記憶されるのを許容する。

【0014】

本発明の別の側面において文字列フォーマッタは上記データベースを検索する前に上記照合対象語句をテキスト・フォーマットする。また単語ウェイトは、上記データベースを検索する前に個々の単語の相対有意性に従い上記照合対象語句内の個々の単語を加重する。これらの工程によれば上記データベースの更に高速かつ正確な検索が許容される。

【0015】

上記自然言語プロセッサ内の検索エンジン、上記一致エントリに対する確信値を生成する。また上記自然言語プロセッサは上記確信値をスレッショルド値と比較する。ブーリアン・テストは、上記照合対象語句からの必要個数の単語が上記一致エントリ内に存在するか否かを決定する。このブーリアン・テストは、上記検索エンジンにより戻される結果を検証する役割を果たす。

【0016】

上記照合対象語句からの必要個数の単語が上記一致エントリ内に存在しなければ、曖昧さを排除すべく上記自然言語プロセッサはユーザに対し、上記一致エントリが上記発語の正しい解釈であるか否かを問い質す。上記一致エントリが上記発語の正しい解釈でなければ、上記自然言語プロセッサはユーザに対し付加的情報を読み質す。この付加的情報により、上記データベース、上記第1文法ファイ

ルおよび上記第2文法ファイルの少なくともひとつが更新される。この様に本発明は付加的な発語の意味を適合的に“学習”することにより、上記ユーザ・インタフェースの効率を高める。

【0017】

上記音声プロセッサは、引き続き発語に対する引き続き照合対象語句に関連付けられた文脈固有文法を有効化して検索する。これにより、最も適切な単語および語句が最初に検索されて音声認識時間を短縮することが確実となる。

【0018】

概略的に本発明は、ウェブ・ページなどのネットワーク・オブジェクトとの音声対話の為にコンピュータを更新する方法を含む。最初に、上記ネットワーク・オブジェクトを上記音声対話システムに関連付けるネットワーク・オブジェクト・テーブルがネットワークを介して上記コンピュータへと転送される。上記ネットワーク・オブジェクト・テーブルの場所は、特定のインターネット・ウェブ・サイトに、または、複数のネットワーク・オブジェクトに対するネットワーク・オブジェクト・テーブルを記憶する統合箇所にて、上記ネットワーク・オブジェクト内に埋め込まれる。上記ネットワーク・オブジェクト・テーブルは、上記ネットワーク・オブジェクトに一致するエン트리に関して検索される。上記ネットワーク・オブジェクトにエントリが一致すると、テキスト音声スピーカから音声発生され、文脈固有文法ファイルが使用され、または、自然言語プロセッサ・データベースが使用されるなどの動作が行われ得る。上記ネットワーク・オブジェクト・テーブルは、ダイアログ定義ファイルの一部とされ得る。ダイアログ定義ファイルは、文脈固有文法ファイルを、または、自然言語プロセッサ・データベースに対するエントリを、または両者を含み得る。

【0019】

本発明の別の側面において、ネットワーク・インタフェースは上記ネットワーク上からダイアログ定義ファイルを転送する。上記ダイアログ定義ファイルはネットワーク・オブジェクト・テーブルを含む。データ・プロセッサは、上記ネットワーク・オブジェクトに一致するテーブル・エントリに関して上記ネットワーク・オブジェクト・テーブルを検索する。この一致テーブル・エントリが発見さ

れたなら、アプリケーション・インタフェースは上記一致エントリにより特定された動作を実施する。

【0020】

本発明の別の側面においては、ネットワークと関連付けられたダイアログ定義ファイルが発見されて読み出される。上記ダイアログ定義ファイルは、ウェブ・サイト、記憶媒体などの種々の場所から、または、複数のネットワーク・オブジェクトに対するダイアログ定義ファイルを記憶する場所から読み出され得る。上記ネットワーク・オブジェクトに一致するテーブル・エントリを発見すべく、上記ダイアログ定義ファイル内に含まれたネットワーク・オブジェクト・テーブルが検索される。上記一致エントリは上記ネットワーク・オブジェクトと関連付けられた動作を定義し、その動作はその後上記システムにより実施される。ネットワーク・オブジェクト・テーブルの他にも、上記ダイアログ定義ファイルは文脈固有文法を、自然言語プロセッサ・データベースに対するエントリを、または両者を含み得る。

【0021】

本発明の特徴、目的および利点は添付図面に關して以下に示された詳細な説明から更に明らかとなるが、同一の参照符号は各図を通して対応している。

(好適実施例の詳細な説明)

次に、図1の好適なコンピュータ・システム100の機能ブロック図を参照して本発明を開示する。図1においてコンピュータ・システム100は、中央処理ユニット(CPU)102を含んでいる。CPU 102は、図3A乃至図3Dに關して本明細書中で記述される方法を実施すべく適切にプログラムされた当業界では公知の任意の汎用マイクロプロセッサもしくはマイクロコントローラとされ得る。上記CPUをプログラムするソフトウェアは記憶媒体108から、又は、コンピュータ・ネットワークを介した別の場所から読み取られ得る。例えばCPU 102は、Intel社などにより製造されたPentium IIプロセッサなどの習用のマイクロプロセッサとされ得る。

【0022】

CPU 102は、ディスプレイ104、手動入力器106、記憶媒体108

、マイクロホン110、スピーカ112、データ入力ポート114およびネットワーク・インタフェース116などの複数の周辺装置と通信する。ディスプレイ104は、ユーザに対してイメージおよびテキストを視覚表示する当業界で公知のCRT、LCD画面、タッチパネルまたは他のモニタなどの視覚的ディスプレイとされ得る。手動入力器106は、習用のキーボード、キーパッド、マウス、トラックボール、または、データの手動入力のために当業界で公知の他の入力デバイスとされ得る。記憶媒体108は、磁気ディスク・ドライブ、フロッピー（登録商標）・ディスク・ドライブ、CD-ROM、シリコン・メモリ、または、データの記憶および読出しのために当業界で公知の他のメモリ・デバイスとされ得る。重要な点として、記憶媒体108はCPU 102から離間して配置され得る。と共に、ローカル・エリア・ネットワーク（LAN）もしくはワイド・エリア・ネットワーク（WAN）もしくはインターネットなどのネットワークを介してCPU 102に接続され得る。マイクロホン110は、CPU 102に対してオーディオ信号を提供する当業界で公知の任意の適切なマイクロホンとされ得る。スピーカ112は、CPU 102からのオーディオ信号を再生する当業界で公知の任意の適切なスピーカとされ得る。マイクロホン110およびスピーカ112は、適切なならばデジタル/アナログおよびアナログ/デジタル変換回路を含む。データ入力ポート114は、RS-232、ユニバーサル・シリアル・バスなどのデータ・プロトコルを使用して外部付属品とインタフェースする当業界で公知の任意のデータ・ポートとされ得る。ネットワーク・インタフェース116は、コンピュータ・ネットワークを介してファイルを通信もしくは転送する当業界で公知の任意のインタフェースとされ得るが、斯かるネットワークの例としては、TCP/IP、イーサネット（登録商標）、またはトークン・リング・ネットワークが挙げられる。これに加え、一定のシステム上にてネットワーク・インタフェース116はデータ入力ポート114に接続されたモデムから成り得る。

【0023】

故に図1は、コンピュータ・システム100の各機能要素を示している。コンピュータ・システム100の各要素は、上述の如き適切な既製の構成要素とされ

得る。本発明は、音声を使用してコンピュータ・システム100との人的対話を行う方法およびシステムを提供する。

【0024】

図8に示された如くコンピュータ・システム100は、コンピュータ・ネットワークの集合であるインターネット700に接続され得る。インターネット700に接続する為に、コンピュータ・システム100はデータ入力ポート114に接続されたモデムであるネットワーク・インタフェース116、または、当業界で公知の他の任意の方法を使用し得る。ウェブ・サイト710、他のコンピュータ720および記憶媒体108もまた、当業界で公知の斯かる方法によりインターネットへと接続され得る。

【0025】

次に図2を参照すると、図2はCPU 102および記憶媒体108の拡大機能ブロック図を示している。図2の各機能要素は全体的にソフトウェアもしくはハードウェアまたは両者にて具現され得ることは理解される。ソフトウェア実施例の場合、そのソフトウェアは記憶媒体108にて、または、コンピュータ・ネットワークを介した代替的箇所にて読み取られ得る。CPU 102は、音声認識プロセッサ200、データ・プロセッサ201、自然言語プロセッサ202およびアプリケーション・インタフェース220を含む。データ・プロセッサ201は、ディスプレイ104、記憶媒体108、マイクロホン110、スピーカ112、データ入力ポート114およびネットワーク・インタフェース116とインタフェースする。データ・プロセッサ201によれば上記CPUは、これらのソースからデータを発見して読み出し得る。自然言語プロセッサ202は更に、変数リプレーサ204、文字列フォーマッタ206、単語ウェイト（word weight）208、ブリーアン・テスト210、代名詞リプレーサ211および検索エンジン213を含んでいる。記憶媒体108は、複数の文脈固有文法ファイル212、一般文法ファイル214、聴音書取文法216、および自然言語処理（NLP）データベース218とを含んでいる。好適実施例において各文法ファイル212、214および216は、ユーザが話す言語の構造を記述するバックス・ナウア記法（BNF）ファイルである。BNFファイルは言語の

構造を記述するものとして当業界で公知ことから、BNFファイルの詳細は本明細書中では論じない。BNFファイルのひとつの利点は、語句もしくは単語のシーケンスを記述する上でこれらの単語シーケンスの全ての組合せを明示的に記述することなく、階層ツリー状構造が使用され得ることである。故に、該好適実施例においてBNFファイルを使用することにより記憶媒体108内の各ファイル212、214および216の物理的サイズは最小化され、以下に記述される如くこれらのファイルが有効化されて検索され得る速度が高まる。但し代替実施例においては、他のファイル構造が使用される。

[0026]

次に図3A乃至図3Dのプロローチャートを参照し、図2の各機能要素の動作および相互作用を記述する。図3Aにおいて、フローはブロック300にて開始し、音声プロセッサ200に対して発語 (utterance) を提供する。発語とは、始まりおよび終わりを有する一連の音であり、話された単語を一個以上含む得る。話された単語を捕捉するマイクrohホン110はブロック300の工程を実施し得る。代替的に、発語はデータ入力ポート114を介して記憶媒体108から音声プロセッサ200へと提供され得る。好適には、上記発語は公知の“wav”オーディオ・ファイル・フォーマットなどのデジタル・フォーマットである。

[0027]

フローは判断302に進み、其処で音声プロセッサ200は、文脈固有文法212のひとつが有効化されているか否かを決定する。もし文脈固有文法212が有効化されていれば、ブロック304にては各文脈固有文法212が検索される。好適実施例において各文脈固有文法212は、起源文脈 (parent context) に関する単語および語句を含むBNFファイルである。概略的に、文脈が対象領域である。例えばパーソナルコンピュータに適用可能な本発明の実施例において、文脈の例は“ニュース (news)”、または“気象 (weather)”もしくは“株 (stocks)”などである。斯かる場合に各文脈固有文法212は、これらの文脈の夫々に対応するコマンド、制御単語、説明、修飾語もしくはパラメータを含むことになる。各文脈を使用すると、情報の形式

に対して階層構造が提供される。各文脈およびそれらの用法は、NLPデータベース218を参照して以下で更に説明する。

[0028]

もし文脈固有文法212が有効化されていれば、文脈固有文法212はブロック300にて提供された上記発語に対する一致に関して検索される。但し文脈固有文法212が未だ有効化されていなければ、フローはブロック308へと進み、其処で一般文法214が有効化される。

[0029]

好適実施例において一般文法214は単語および語句を含むBNFファイルであり、斯かる単語および語句自体は起源文脈に属しないが、文脈固有文法ファイル212が存在する関連文脈を有し得る。換言すると、一般文法214中の単語および語句は階層的な文脈構造のルートに存在し得る。例えばパーソナルコンピュータに適用可能な一実施例において、一般文法214はコマンドおよび制御語句を含むことになる。

[0030]

ブロック310にて一般文法214は、ブロック300にて提供された発語に対する照合対象単語もしくは語句に関して検索される。ブロック312にては、一致が見られたか否かに依存して判定が為される。もし一致が見られなければ、ブロック314にて聴音書取文法216が有効化される。好適実施例において聴音書取文法216は各単語のリストを含むBNFファイルであり、各単語自体は起源文脈も関連文脈も有さない。例えばパーソナルコンピュータに適用可能な一実施例において、聴音書取文法216は一般的辞書に類似した一般的単語の比較的に大きなリストを含んでいる。

[0031]

ブロック316にて上記聴音書取文法は、ブロック300にて提供された発語の各単語に対する照合対象単語に関して検索される。判定ブロック318にて照合対象単語が発見されなければ、ブロック320にてはディスプレイ104に視覚的エラー・メッセージが選択的に表示され、または、スピーカ112により可聴エラー・メッセージが選択的に再生される。而して上記プロセスは、ブロック

300にて音声プロセッサ200に対し別の発語が提供されるまで停止する。

【0032】

故に上記記述から理解され得る如く、音声プロセッサ200に対して発語が提供されたとき、もし有効化された文脈固有文法ファイル212があれば、その文脈固有文法ファイル212が先ず検索される。有効化された文脈固有文法ファイル212内で一致が無ければ、一般文法214が有効化され且つ検索される。もし一般文法214でも一致しなければ、聴音書取文法316が有効化され且つ検索される。

【0033】

好適実施例において音声認識プロセッサ200が文脈固有文法212または一般文法214のいずれかを検索しているとき、それは“コマンド/制御”モードに在ると称される。このモードにおいて音声認識プロセッサ200は、発語全体を全体として上記文法中の各エントリと比較する。これと対照的に、音声認識プロセッサ200が上記聴音書取文法を検索しているとき、それは“聴音書取”モードに在ると称される。このモードにおいて音声認識プロセッサ200は、聴音書取文法216内の各エントリに対し、上記発語を一度に一単語ずつ比較する。上記コマンド/制御モードにおいて発語全体に対する一致を検索することは、聴音書取モードにおいて一度に一単語ずつ検索することよりも一般的に高速であると期待される。

【0034】

更に、任意の個々の文脈固有文法212は一般文法214よりもサイズが小さく(すなわち合計の単語および語句が少なく)、一般文法214自体は聴音書取文法216よりもサイズが小さいことも期待される。故に、有効化された任意の文脈固有文法212を最初に検索すれば、少なくとも部分的な理由としてファイル・サイズが小さい故に、一致があれば更に迅速に発見されると思われる。同様に、聴音書取文法216の前に一般文法214を検索することにより、一致があれば更に迅速に発見されると思われる。

【0035】

また、文脈固有文法212および一般文法214の両者に対して適合的に付加

される本発明の機能に関して以下で更に説明される如く、これらの文法は最も一般的な発語を含んでいる。その故に、聴音書取文法216におけるよりも文脈固有文法212もしくは一般文法214における方が更に迅速に一致が発見されることが期待される。

【0036】

最後に、以下で更に説明される如く、有効化された文脈固有文法212における単語および語句はユーザにより発語される可能性が更に高いものである、と言うのも、それらの単語および語句はユーザが最も最近に話した主題に関連する単語だからである。これにより、有効化された文脈固有文法212に従ってユーザの各単語の意味が解釈され乍ら、ユーザは断片的な文を使用して更に会話的な形式で話すことができる。

【0037】

本発明は上述のシーケンスで検索を行うことから、予測される全ての単語および語句を含む単一の大きなリストで一度にひとつのエントリを検索するよりも更に効率的な検索を行い得る。

【0038】

判定306に戻ると、文脈固有文法212内で一致が発見されたなら、図3Aに示された3つの可能な次工程が在る。有効化された文脈固有文法212内の各一致エントリに対し、音声認識プロセッサ200により行われる関連動作が在り得る。ブロック322はひとつの動作がアプリケーション・インタフェース220に指示を行い、該アプリケーション・インタフェース220が別体のソフトウェア・アプリケーションもしくはエントリに関して一定の動作を行い得ることを示している。例えばアプリケーション・インタフェース220は、Microsoftによる音声アプリケーション・プログラミング・インタフェース(SAPI)規格を使用して外部アプリケーションと通信し得る。上記外部アプリケーションに対しては例えば、特定のインターネット・ウェブ・サイトURLにアクセスすること、または、テキストを音声に変換して特定語句を話すこと、が指示され得る。他の動作は、図4のNLPデータベース218に関して以下で更に論じられる如く行われ得る。

【0039】

ブロック324は別の動作として、自然言語処理(NLP)データベース218におけるひとつの行(row)に直接的にアクセスすることにより、以下で更に記述される各自然言語処理工程をバイパスし得ることを示している。ブロック326は別の動作として、文脈固有文法306で発見された照合対象単語もしくは語句に対し、有効化された文脈の単語もしくは語句を前方付属(prepend)し得ることを示している。例えば有効化された文脈が“映画(movies)”であり且つ照合対象発語が“8時(8 o'clock).”であれば、単語“映画”が語句“8時”に前方付属されて“8時の映画(movies at 8 o'clock)”という語句が形成される。

【0040】

同様に、もし一般文法214内に一致が発見されればフローは、ブロック32に進み其処でアプリケーション・インタフェース220に上記の動作を行うことを指示し、または、ブロック324に進み其処でNLPデータベース内のひとつの行が直接的にアクセスされる。但し一般文法214において一致が発見されたとときに文脈の前方付属は生じない、と言うのも、上述の如く一般文法214内の各エントリはそれら自体が起源文脈を有さないからである。

【0041】

代替的に、ブロック301にては手動入力された単語が捕捉され、上記自然言語プロセッサへと入力される。

最後に図3Aを参照すると、各単語は手動入力器106を介して手動で入力され得る。この場合に音声認識は不要であるが、入力された単語の自然言語処理は依然として望ましい。故にフローは図3Bへと進む。

【0042】

図3Bのブロック328において、自然言語プロセッサ202は自然言語処理分析の為に上記語句をフォーマットする。このフォーマットは、文字列フォーマット206により達成されると共に、各単語間の二重空白を除去し、全ての文字を小文字(もしくは大文字)にし、短縮を広げる(例えば、“its”を“itis”に変更する)などのテキスト処理を含み得る。このフォーマット工程

の目的は、語句を解析に対して準備することである。

【0043】

フローはブロック330へと続き、該ブロックにては、NLPデータベース218へのアクセスに備えて変数リプレーサ204により関連ワイルドカード機能により一定の“単語変数(word-variable)”が置換される。本明細書中で使用される“単語変数”とは、量、日付、時間、通貨などを表す単語もしくは語句を指している。例えば一実施例において“8時にはどの映画が放映されるか(what movies are playing at 8 o'clock)”という語句はブロック330にて“時間にはどの映画が放映されるか(what movies are playing at time)”と変換されるが、この場合に“時間(time)”は任意の時間値を表すべく使用されるワイルドカード機能である。別の場合の一実施例においては、“IBMの株を100ドルで売る(sell IBM stock at 100 dollars)”という語句はブロック330にて“IBMの株をドルで売る(sell IBM stock at dollars)”と変換されるが、この場合に“ドル(dollars)”は任意のドル値を表すべく使用されるワイルドカード機能である。この工程は簡単なループにより達成され得るが、該ループは“ドル”もしくは“時(o'clock)”などの単語であるキー・トークンに関して語句を検索し、且つ、特定のワイルドカード機能により単語変数を置換するものである。置換が行われた語句内の箇所を追跡する為に配列が使用される。これにより、NLPデータベース218が検索された後で一定の位置にて上記語句へと元の単語変数を戻すべく再置換することが許容される。

【0044】

ブロック330にて単語変数を関連ワイルドカード機能により置換する目的は、NLPデータベース218内に存在すべきエントリの個数を減少することである。例えばNLPデータベース218は、8時(8 o'clock)、9時(9 o'clock)、10時(10 o'clock)などの別体のエントリではなく、“時間にはどの映画が放映されるか(what movies are playing at time)”という語句のみを含む。NLPデータ

ベース218は以下で更に記述される。

【0045】

ブロック332にては、上記語句内の各代名詞が代名詞リプレサ211により適切な名称で置換される。例えば一実施例において、代名詞“私(I)”、“私の(my)”もしくは“私のもの(mine)”は話し手の氏名で置換される。この工程の目的は、ユーザがNLPデータベース218内にユーザ特有の事実を記憶してアクセスするのを許容することである。例えば“誰が私の子供ですか(who are my children)”という文は、“誰がDeanの子供ですか(who are Dean's children)”に変換される。この場合に“Dean”は話し手の固有名詞である。此処でも、この工程は簡単なループにより達成され得るが、該ループは、代名詞に関して上記語句を検索し、発見された代名詞を適切な固有名詞に置換する。置換が行われた語句内の箇所を追跡する為に配列が使用され得る。

【0046】

ブロック334にては語句内の個々の単語が単語ウエイタ208により、該語句の全体的意味に対する“重要性”もしくは“有意性”に従い加重される。例えば一実施例においては、割当てられる3種の加重係数が在る。最低加重係数は、“a”、“an”、“the”および他の品詞などの単語に割当てられる。最高加重係数は、上記語句の意味に対して有意性を有すると思われる単語に対して与えられる。例えばそれらの単語とは、NLPデータベース218内の全ての動詞、名詞、形容詞および固有名詞である。中間加重係数は、上記語句内の他の全ての単語に与えられる。この加重の目的は、NLPデータベース218に対する更に強力な検索を許容することである。

【0047】

図4には、一実施例におけるNLPデータベース218の選択列の例が示されている。NLPデータベース218は複数の列400〜410および複数の行412A〜412Nを備えている。列400において各エントリは、上記NLPデータベースに対して“既知”の語句を表している。列402においては、列400における各エントリに対して必要な単語の個数が示される。列404において

は、列400における各エントリに対する関連文脈または下位文脈が示される。列408および410においては、列400における各エントリに対し、一個以上の関連動作が示される。尚、図4に示されたNLPデータベース218は本発明を教示すべく簡略化された例に過ぎないことを銘記されたい。他の実施例は、種々のエントリにより更に多くのもしくは更に少ない列を有し得る。

【0048】

図3Bに戻ると、ブロック336にてNLPデータベース218は、該NLPデータベース218の列400内のエントリが上記語句内の各単語（またはそれらの同義語）のいずれかを含むか否かということと、これらの単語の相対重みとに基づき、上記語句に対する可能な一致に関して検索される。ブロック338にては、上記語句内における各単語の発生回数と、各単語の相対重みに基づき、可能な一致エントリに対して確信値が生成される。データベースの加重単語検索は当業界で公知であり、且つ、バージニア州、アーリントン郡のDT Software, Inc. による製品“dtsearch”などの市販の検索エンジンにより実施され得る。同様に、同義語を使用した検索は、当業界で公知であり、且つ、ニュージャージー、プリンストンのプリンストン大学の認知科学研究所(Cognitive Science Laboratory)により開発された“WordNet”の如き公開ツールなどを使用して達成され得る。上記検索エンジンは、自然言語プロセッサ202の一体的部分とされ得る。

【0049】

判定340にて自然言語プロセッサ202は、可能な一致エントリのいずれかが、一定の所定最小スレッショルドT以上である確信値を有するか否かを決定する。このスレッショルドTは、NLPデータベース218内の各エントリのいずれかに対して上記語句が一致したか否かに関する判定が為される為の容認可能最低確信値を表す。もしT以上の確信値を有する可能な一致エントリが無ければローはブロック342へと進み、選択的エラー・メッセージがユーザに対してディスプレイ104上で視覚的表示され、または、スピーカ112を介して音響的に再生される。一実施例においてユーザに対して表示されるエラー・メッセージがあれば、そのエラー・メッセージの形式は、最高確信NLPデータベース・エ

ントリ内で何個の“適合 (hit)” (すなわち、上記語句の何個の照合対象単語) が発見されたかに依存し得る。また、ゼロもしくは1個の適合であった場合には、2個以上の適合だった場合とは異なる形式のエラー・メッセージが生成される。

【0050】

但しNLPデータベース218内において、当該エントリに対してT以上の確信値が存在する少なくとも一個のエントリが在れば、フローはブロック344へ進み其処で上記語句からは“ノイズ (noise)”単語が廃棄される。“ノイズ”単語としては、上記語句内の他の単語と比較して該語句の全体的意味に関連してそれほど寄与しない単語が挙げられる。斯かる単語としては、冠詞、代名詞、接続詞、および同様の性質の単語が挙げられる。“非ノイズ (non-noise)”単語は、上記語句の全体的意味に関して相当に寄与する単語が挙げられる。“非ノイズ”単語としては、動詞、名詞、形容詞、固有名称、および同様の性質の単語が挙げられる。

【0051】

フローは図3Cに進みブロック346にては、最高確信一致エントリに関してNLPデータベース218の列402から非ノイズ単語要件が読出される。例えば、もし最高確信照合対象語句が行412A内のエントリ (例えば、“時間にはどの映画が放映されるか (what movies are playing at time)”) であれば、非ノイズ単語必要個数は3である。

【0052】

判定348にては、上記語句からの非ノイズ単語必要個数が、NLPデータベース218から読出された最高確信エントリ内に実際に存在するか否かを決定するテストが行われる。このテストはブロック336において実施された適合形式検索 (relevance-style search) の精度の検証であり、上記語句と容認可能に一致しなくてもエントリは最小スレッショルドTよりも大きな確信値を生成する可能性が在ることは理解される。

【0053】

判定348で実施される上記テストの性質は、ブーリアン・テスト210によ

り実施される論理的“AND”テストである。このテストは、上記語句 (もしくはその同義語) 内の非ノイズ単語の各々が上記最高確信エントリ内に実際に存在するか否かを決定するものである。もし上記最高確信エントリ内に実際に存在する十分な個数の必要単語が在れば、フローはブロック350へと進み、其処で自然言語プロセス202はアプリケーション・インタフェース220に対して列408もしくは410からの関連動作を行うことを指示する。付加的な動作列も存在し得ることは理解されよう。

【0054】

例えば、最高確信エントリが行412A内のエントリであり、且つ、判定348のブーリアン・テストが、列400内のエントリ内の語句には実際に3個の非ノイズ単語が在ることを決定したなら、列408における関連動作 (例えば、映画ウェブ・サイトにアクセスする) が行われる。上記NLPデータベースにおける他の各エントリは他の関連動作を有する。例えば、最高確信エントリが行412Eにおけるエントリ (例えば、“いま何時 (what time is it)”) であれば関連動作は、自然言語プロセス202が (不図示の) テキスト/音声アプリケーションに対し、スピーカ112を介してユーザに現在時刻を音声発生することを指示することである。他の例として、もし最高確信エントリが行412Nのエントリ (例えば、“ニュースを見せて下さい (show me the news)”) であれば、第1の関連動作はインターネット上の所定のニュース・ウェブ・サイトにアクセスすることであり、且つ、第2の関連動作は (不図示の) イメージ表示アプリケーションに対し、上記ニュースに関連するイメージを表示することを指示することであり得る。別のもしくは付加的な動作も実施され得る。

【0055】

同様に、もし最高確信エントリが判定348で決定された如き上記語句からの非ノイズ単語必要個数を含むならば、自然言語プロセス202は音声認識プロセス200に対し、列404の関連文脈に対して文脈固有文法212を有効化することを命令する。故に行412Aに対しては、“映画 (movies)”という文脈に関する文脈固有文法212が有効化される。故に、図3Aのブロック

300にて音声認識プロセッサ200に対して次の発語が提供されたとき、音声認識プロセッサ200は一般文法214を検索する前に“映画 (movies)”に対する有効な文脈固有文法212を検索する。前述の如く、適切な文脈固有文法212を有効化すると高速で好結果な音声認識の可能性が高まると共に、会話的な形式でコンピュータと対話するユーザの能力が強化される。

[0056]

但し判定348にて、NLPデータベース218から読出された最高確信エンタリ内に上記語句からの非ノイズ単語必要個数が実際に存在しなければ、フローはブロック354に進み、其処でユーザはディスプレイ104もしくはスピーカ112を介し、最高確信エンタリが表されたか否かが問われる。例えばユーザが“今日のIBM株の売りはいくらか (How much is IBM stock selling for today)”と発語すると、NLPデータベース218内の最高確信エンタリは行412B内のエンタリである。この場合、適合率が高いが、必要単語（もしくはそれらの同義語）の個数は十分でない。故にユーザはブロック354にて、ユーザが“1998年8月28日のIBM株の値段はいくらか (what is the price of IBM stock on August 28, 1998)”を意味したか否かが問われる。

[0057]

ユーザは、肯定的にもしくは否定的に応答し得る。ユーザが肯定的に応答したと判定356で決定されたなら、ブロック350にては最高確信エンタリに関連する動作が行われ、且つ、ブロック352にては関連する文脈固有文法212が有効化される。

[0058]

但し判定356にてユーザが否定的に応答したならフローは図3Dへと進み、其処で、上記最高確信エンタリに関してNLPデータベース218の列404から関連文脈が読出され、且つ、ブロック360にてユーザは文脈用対話型ダイアログを使用して情報が問われる。例えばユーザが“XICOR株の今日の値段はいくらか (what is the price of XICOR stock

today)”と発語すると共にNLPデータベース218からの最高確信エンタリは行412B（例えば“日付のIBM株の値段はいくらか (what is the price of IBM stock on date)”）であったとすれば、ブロック354でユーザは、それがユーザの意図した処であるか否かが問われる。

[0059]

もしユーザが否定的に応答したなら、“株 (stock)”という文脈がブロック358にて列404から読出され、ディスプレイ104およびスピーカ112を介してユーザには株文脈に対する文脈用対話型ダイアログが呈示される。斯かる文脈用対話型ダイアログは、XICOR株に対する名称および相場表示器記号に関してユーザに問う段階を含み得る。ユーザは必要情報を話すことで応答し得る。各可能的文脈に対しては、異なる文脈用対話型ダイアログが使用され得る。例えば“気象 (weather)”文脈用対話型ダイアログは、気象情報が所望される場所（例えば市）の名称をユーザに問う段階を含み得る。また“ニュース (news)”文脈用対話型ダイアログは、記事の種類、ニュース・ソース、ニュース・サイトに対するインターネットURL、または、他の関連情報をユーザに問う段階を含み得る。

[0060]

上記文脈用対話型ダイアログが完了したなら、NLPデータベース218、一般文法214および文脈固有文法212はブロック362にて新情報を含むべく更新される。この様にして、ユーザは次回にてその情報を要求し、適切な一致が発見され、且つ、更なる情報をユーザに問うこと無く適切な動作が行われる。故に本発明は、ユーザにより発語された語句を認識すべく適格的に“学習”する。

[0061]

本発明の一実施例においては、NLPデータベース218、文脈固有文法212、一般文法214および聴音書取文法216のひとつ以上が、各エンタリに関連する（不図示の）タイムスタンプ値も含み得る。一致エンタリが使用される毎に、そのエンタリに関連するタイムスタンプ値は更新される。定期的間隔にて、または、ユーザにより開始されたとき、一定の日付および時刻より以前のタイム

スタンプ値を有する各エントリは夫々のデータベース／文法から削除される。この様にして上記データベース／文法は古いもしくは期限経過エントリを“一掃”することで効率的サイズに保持され得る。これは、誤った一致の回避も促進する。

【0062】

本発明の一実施例において音声認識および自然言語処理は、文脈依存音声式インタフェースを有するワールド・ワイド・ウェブ（“WWW”もしくは“ウェブ”）などのネットワーク・オブジェクトと対話すべく使用され得る。

【0063】

図5は好適なダイアログ定義ファイル（DDF）500を示すが、該ファイル500は、テキストもしくはグラフィック・ファイル、または、好適実施例にてウェブ・ページなどであるインターネット・オブジェクトに対して上記音声認識および自然言語処理を関連させる為に必要な情報を表している。最も簡単な実施例においてダイアログ定義ファイル500はネットワーク・オブジェクト・テーブル510から成るが、該DDFは図5に示された如く付加的な文脈固有文法ファイル214および付加的な自然言語処理（NLP）データベース218も含み得る。DDF500の好適実施例は、ネットワーク・オブジェクト・テーブル510と、文脈固有文法ファイル214と、自然言語処理データベース218に対するエントリを含むファイルとを含む。これらの構成要素は圧縮されると共に、Lempel-Ziv圧縮などの当業界で公知の任意の方法によりDDFファイル500へと結合され得る。文脈固有文法ファイル214および自然言語処理データベース218は、上述の各段落で説明した。ネットワーク・オブジェクト・テーブル510は、リソースのアドレスをNLPデータベース218内の種々の動作、文法もしくはエントリに対して関連付けるメモリ・ツリー、チェーンもしくはテーブルなどのメモリ構造である。

【0064】

ネットワーク・オブジェクト・テーブル510の好適実施例は図6に示される。図6はメモリ・テーブルを示すが、該メモリ・テーブルは次のものに対するエントリ列を含み得る：ネットワーク・オブジェクト520、テキスト→音声（T

スタンプ値を有する各エントリは夫々のデータベース／文法から削除される。この様にして上記データベース／文法は古いもしくは期限経過エントリを“一掃”することで効率的サイズに保持され得る。これは、誤った一致の回避も促進する。

【0062】

本発明の一実施例において音声認識および自然言語処理は、文脈依存音声式インタフェースを有するワールド・ワイド・ウェブ（“WWW”もしくは“ウェブ”）などのネットワーク・オブジェクトと対話すべく使用され得る。

【0063】

図5は好適なダイアログ定義ファイル（DDF）500を示すが、該ファイル500は、テキストもしくはグラフィック・ファイル、または、好適実施例にてウェブ・ページなどであるインターネット・オブジェクトに対して上記音声認識および自然言語処理を関連させる為に必要な情報を表している。最も簡単な実施例においてダイアログ定義ファイル500はネットワーク・オブジェクト・テーブル510から成るが、該DDFは図5に示された如く付加的な文脈固有文法ファイル214および付加的な自然言語処理（NLP）データベース218も含み得る。DDF500の好適実施例は、ネットワーク・オブジェクト・テーブル510と、文脈固有文法ファイル214と、自然言語処理データベース218に対するエントリを含むファイルとを含む。これらの構成要素は圧縮されると共に、Lempel-Ziv圧縮などの当業界で公知の任意の方法によりDDFファイル500へと結合され得る。文脈固有文法ファイル214および自然言語処理データベース218は、上述の各段落で説明した。ネットワーク・オブジェクト・テーブル510は、リソースのアドレスをNLPデータベース218内の種々の動作、文法もしくはエントリに対して関連付けるメモリ・ツリー、チェーンもしくはテーブルなどのメモリ構造である。

【0064】

ネットワーク・オブジェクト・テーブル510の好適実施例は図6に示される。図6はメモリ・テーブルを示すが、該メモリ・テーブルは次のものに対するエントリ列を含み得る：ネットワーク・オブジェクト520、テキスト→音声（T

214で置換され且つ既存ファイルは無効化されることが示される。“文法追加”フラグ528は、既存の文脈固有文法ファイルと同時に新たな文脈固有文法ファイルが有効化されるべきことを示す。

【0067】

最後に、残りの列のエントリは、“はい/いいえ”文法構造に關している。もし“はい/いいえ”フラグ530がマークされたなら、標準的な“はい/いいえ”文法が有効化される。標準的な“はい/いいえ”文法が有効化されたとき、コンピュータに対して肯定的なコマンドを話すと、コンピュータは“はいなら行う”エントリ532内に示されたコマンドを実行する。同様に、コンピュータに対して否定的なコマンドを話すと、コンピュータは“いいえなら行う”エントリ534内に示されたコマンドを実行する。“はいなら行う”列532および“いいえなら行う”列534内の各エントリは、コマンドと、NLPデータベース218内に記憶されたコマンドに対するポインタと、のいずれかとされ得る。例えば、行540Bに示された如く、“はい/いいえ”フラグがマークされる。コンピュータに対して“はい”などの肯定的回答が与えられ、と、“はいなら行う”エントリ532内の対応コマンドが実行され；この特定の場合にエントリは数字“210”であり、これは上記NLPデータベース内の第210番目のコマンドへの参照である。“いいえ”という答えならコンピュータは上記NLPデータベース内の第211番目のコマンドを実行する。

【0068】

図7Aに行くと、ワールド・ワイド・ウェブをブラウズするコンピュータなどのインターネットワークを行うコンピュータに対して音声および音響でコマンドを提供する方法およびシステムが示されている。図7A乃至図7Cの方法は、図3A乃至図3Dの方法と組合せて使用され得る。ブロック602にて、ワールド・ワイド・ウェブ・ブラウザに対してはウェブ・サイトURL（ネットワーク・オブジェクト）が提供される。上記ウェブ・ブラウザはインターネット上をナビゲートすべく使用されるプログラムであり、当業界で公知である。ブロック602においてブラウザに対してURLを提供する段階は、ユーザがURLを手動タイプ入力するか、または、選択されたウェブ・サイトURLに対する“リンク”

をユーザに選択させる如く可及的に容易である。この段階はまた、NLPデータベース218内の各エントリに關連付けられた動作に關して前述された発声コマンドの結果でもあり得る。URLが与えられたならコンピュータはブロック604にて、上記URL内で指定されたウェブ・サイトのインターネット・アドレスを決定し得るか否かを判定せねばならない。この決定プロセスは、当業界で公知のプロセスである。もしコンピュータが上記インターネット・アドレスを得なければ、ブロック605にてはブラウザ画面内にエラー・メッセージが表示され、システムはその初期開始状態600へと戻る。もし上記インターネット・アドレスが決定されるなら、コンピュータはブロック606にてウェブ・ページに対する要求をウェブ・サイトへと送信する。

【0069】

ブロック608にては、上記ウェブ・サイトがウェブ・ページを送信するか否かに依存して判定が為される。もし上記ウェブ・サイトが応答しなければ、または、ウェブ・ページの送信が不首尾となれば、ブロック605にてはブラウザ画面内にエラー・メッセージが表示され、システムはその初期開始状態600へと戻る。もし上記ウェブ・サイトがウェブ・ページを戻したなら、ブロック610ではブラウザ画面内にそのウェブ・ページが表示される。

【0070】

判定ブロック612にてコンピュータ100は、上記ウェブ・サイトに対応するDDFファイル500が既に該コンピュータ100内に存在するか否かを決定する。もし上記DDFファイルが存在すれば、フローは図7Cへと進み、存在しなければフローは図7Bへと進む。

【0071】

図7Bに行くと、もしDDFファイル500が存在しなければ上記コンピュータは、DDFファイル500の場所がURLとしてウェブ・ページのハイパー・テキスト・マークアップ言語（HTML）内にコード化されているか否かを吟味する。（HTMLは当業界で公知であることから、該言語の詳細は本明細書では論じない。）HTMLコード内へのDDFファイル場所のコード化は、

```
<meta DDF="http://www.conversationa
```

```
l s y s . c o m / C o n v e r s e l t . d d f " >
```

などの様に初期HTMLメタ・タグ内にDDFファイル場所を列挙するか、ブラウザによりサポートされるHTMLの変種へとスクリプト・タグを直接的に書き込むことで行われる：

```
<!--
<DDF="http://www.conversationalsys
.com/ConverseIt.ddf">
```

```
-->
```

もしDDFファイル場所の情報がウェブ・ページ内にコード化されていれば、その場所のインターネット・アドレスはブロック616にて決定され、ブロック626にてコンピュータはDDFファイル500の転送を要求する。

[0072]

代替的に、もしDDFファイル500の場所が上記ウェブ・ページ内にコード化されていないければ、DDFファイルの場所が記憶され得る幾つかの代替的場所がある。DDFファイルの場所は、ルート・ディレクトリ内の一定のファイル箇所などのウェブ・サイトにおいて、または、別のインターネット・サーバもしくは図1の記憶媒体108などの異なる中央箇所において、事前定義箇所内に記憶され得る。ブロック618および620はこれらの可能性に対してテストを行う。ブロック618は、DDFファイルが上記ウェブ・サイトに配置されているかを決定する。この工程にて上記コンピュータは、DDFファイル500の存在に関して照会を行うクエリを上記ウェブ・サイトに送信する。もし上記ウェブ・サイトにDDFファイル500が存在すれば、上記コンピュータはブロック626にてDDFファイル500の転送を要求する。もしDDFファイル500が上記ウェブ・サイトに配置されていないければ、上記コンピュータはブロック620にて上記ウェブ・サイトに関するDDFファイルの存在に関して中央箇所に照会する。もし上記DDFファイルが上記ウェブ・サイトに存在すれば、上記コンピュータはブロック626にて上記DDFファイルの転送を要求する。もしDDFファイル500が発見されなければ、ブロック622にては、ネットワーク・オブジェクト・テーブル510、上記ウェブ・サイトに関連するNLPデータベース

ース218、および、以前に訪れた一切のウェブ・サイトに対する文脈固有文法214、などの既存DDFファイルの既存構成要素が作動停止される。更にブロック624にて、上記ウェブ・サイトは非音声起動ウェブ・サイトとして処理されると共に、標準的文法ファイルのみが使用される。標準的文法ファイルは上記システム上に存在する文法ファイルであるが、上記ネットワーク・オブジェクトに関連する文脈固有文法ファイルと関連する一切の文法は除かれる。

[0073]

もしブロック626にてDDFファイル500が要求され乍らもその転送が不首尾であれば、ブロック622では存在する一切のDDFファイル500の一切の既存構成要素が作動停止され、且つ、ブロック624では上記ウェブ・サイトが非音声起動ウェブ・サイトとして処理され且つ標準的文法ファイルのみが使用される。

[0074]

もしブロック626にてDDFファイル500が要求されると共にその転送が好結果であれば、そのDDFファイルはブロック630にて一切の先行DDFファイルに置き換わる。ブロック632にては、ネットワーク・オブジェクト・テーブル510、文脈固有文法ファイル214およびNLPデータベース218などのDDFファイル500の一切の構成要素が抽出される。図2の機能要素を含め図3A乃至図3Dに示された方法を実現する為に必要なソフトウェアを獲得する為に、同様の技術が使用され得る。

[0075]

フローは図7Cへと移動する。ブロック634にてネットワーク・オブジェクト・テーブル510は上記コンピュータによりメモリ内に読取られる。ブロック636により決定された如く、もし上記ウェブ・ページURLが上記サイト・ネットワーク・オブジェクト・テーブル510内に存在すれば、斯かるウェブ・ページURLは図6に示された如く上記テーブルの行540A乃至540Eにより表される。もし上記ウェブ・ページに対応する行が存在しなければ、上記ウェブ・ページに対する非音声相互作用が存在し、処理は終了する。

[0076]

ブロック636で決定された如くもし上記サイト・ネットワーク・オブジェクト・テーブル510内にウェブ・ページURLが存在すれば、ブロック638にて上記コンピュータは、TTSフラグ522がマークされているか否かをチェックすることにより上記ウェブ・ページにテキスト音声524が関連付けられるか否かを決定する。もしテキスト音声524が在れば、それはブロック640にて音声発生され、フローは継続する。もし判定ブロック642にて決定された如く上記ウェブ・ページに関連付けられた文脈固有文法ファイルが在れば、それはブロック644で有効化され、次にブロック646でNLPデータベース218が有効化される。もし上記ウェブ・ページに文脈固有文法ファイルが関連付けられていなければ、ブロック646にてNLPデータベース218のみが有効化される。上記NLPデータベースが有効化646されたなら、上記システムは上述の如く図3A乃至図3Cの様挙動する。

[0077]

要約すると、本発明はコンピュータのネットワーク用対話型ユーザ・インタフェースに対する方法およびシステムを提供する。ダイアログ定義ファイルを紹介してインターネット・オブジェクトに連結された文脈固有文法を使用することにより、本発明は音声認識時間を短縮すると共に、ユーザが会話的な形式でウェブ・ページなどのインターネット・オブジェクトと対話する能力を高める。更に、種々の文法およびNLPデータベースを適合更新することにより、本発明は対話効率を更に高める。

[0078]

各好適実施例に関する上記記述は、当業者が本発明を実施もしくは使用し得る様に提供されたものである。当業者であればこれらの実施例の種々の改変例は容易に明らかであると共に、本明細書中で定義された包括的原理は発明能力を使用せずとも他の実施例に適用され得る。故に本発明は本明細書中に示された各実施例に限定されることは意図されず、本明細書中に開示された原理および新規特徴に一致する最大範囲に従うものとする。

【図面の簡単な説明】

【図1】本発明と共に使用される好適なコンピュータ・システムの機能プロ

ック図。

【図2】図1のコンピュータ・システムにおける本発明のCPU 102および記憶媒体108の拡大機能ブロック図。

【図3A】コンピュータに対する対話型音声認識／自然言語処理を提供する方法を示すフローチャート。

【図3B】コンピュータに対する対話型音声認識／自然言語処理を提供する方法を示すフローチャート。

【図3C】コンピュータに対する対話型音声認識／自然言語処理を提供する方法のフローチャート。

【図3D】コンピュータに対する対話型音声認識／自然言語処理を提供する方法を示すフローチャート。

【図4】本発明の好適な自然言語処理(NLP)データベースの選択列を示す図。

【図5】本発明に係る好適なデータベース定義ファイル(DDF)を示す図。

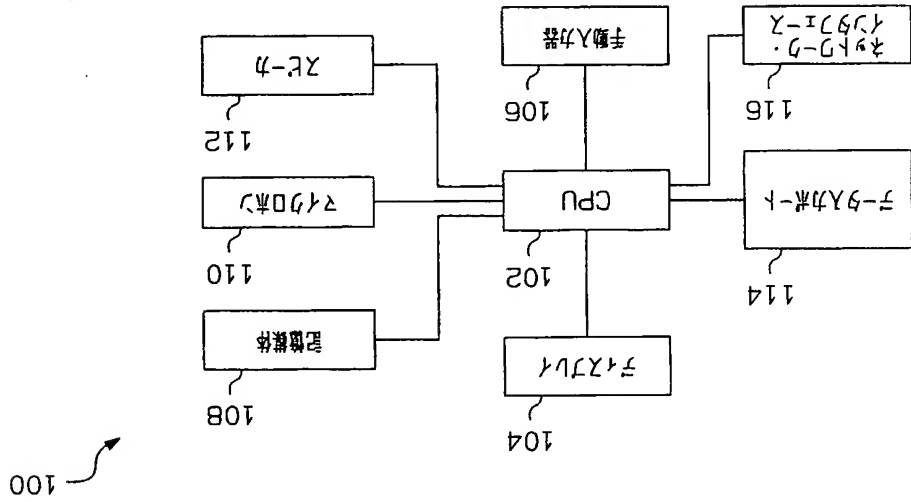
【図6】本発明の好適なネットワーク・オブジェクト・テーブルの選択列を示す図。

【図7A】ウェブ・ページなどのネットワーク・オブジェクトに対する対話型音声認識／自然言語処理のリンクを示す、本発明の方法のフローチャート。

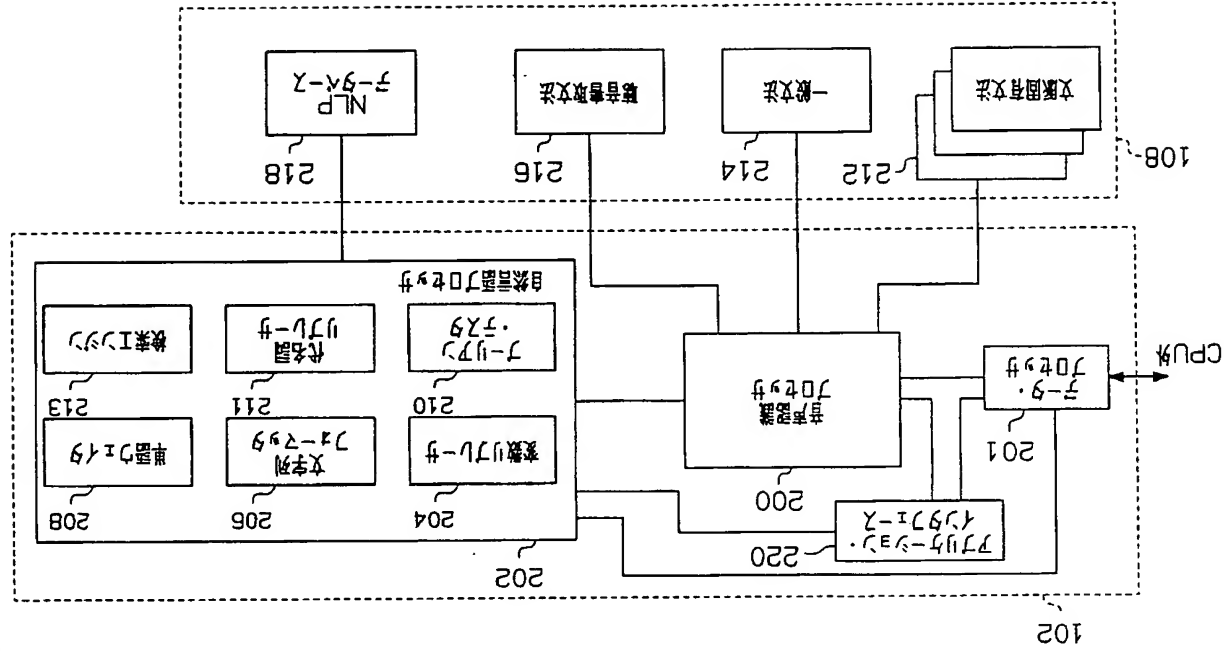
【図7B】ウェブ・ページなどのネットワーク・オブジェクトに対する対話型音声認識／自然言語処理のリンクを示す、本発明の方法のフローチャート。

【図7C】ウェブ・ページなどのネットワーク・オブジェクトに対する対話型音声認識／自然言語処理のリンクを示す、本発明の方法のフローチャート。

【図8】インターネットを介して他のコンピュータ、記憶媒体およびウェブ・サイトに接続しているコンピュータ・システムを示す図。

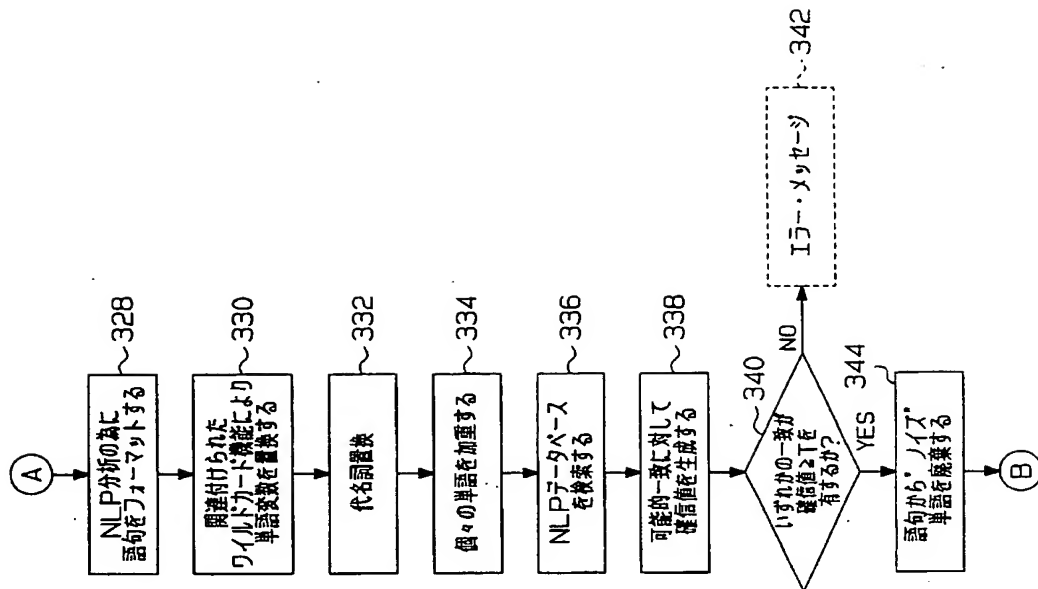


【図1】

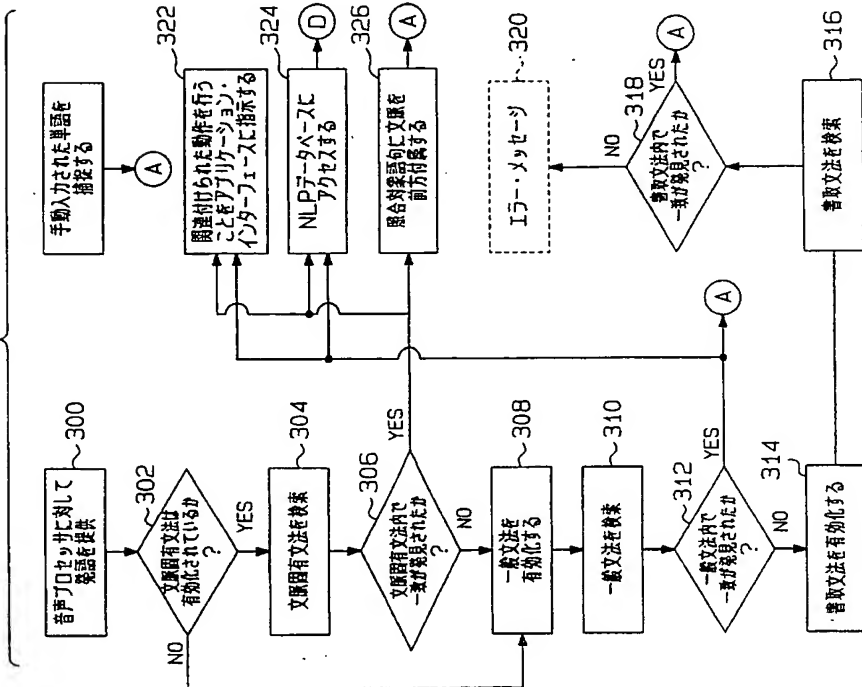


【図2】

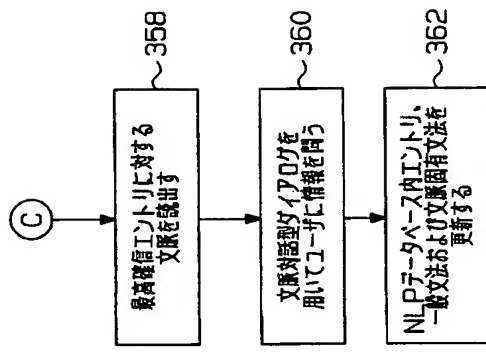
【図 3 B】



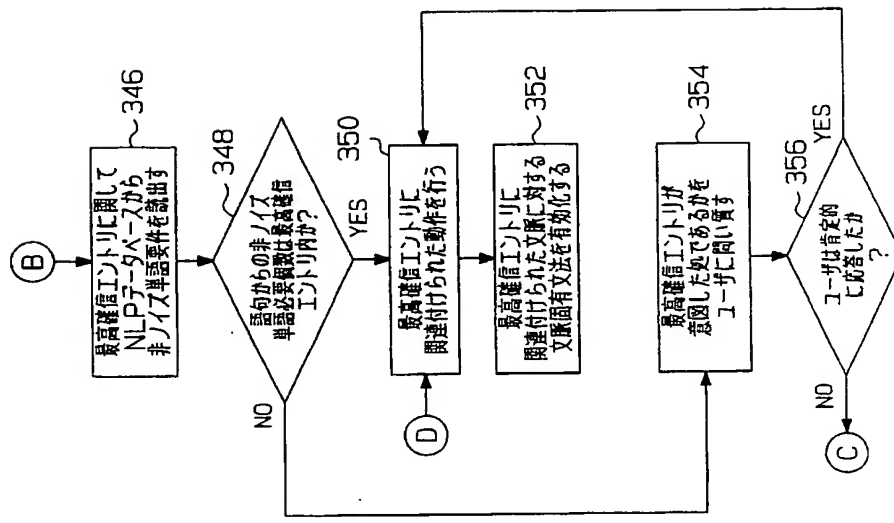
【図 3 A】



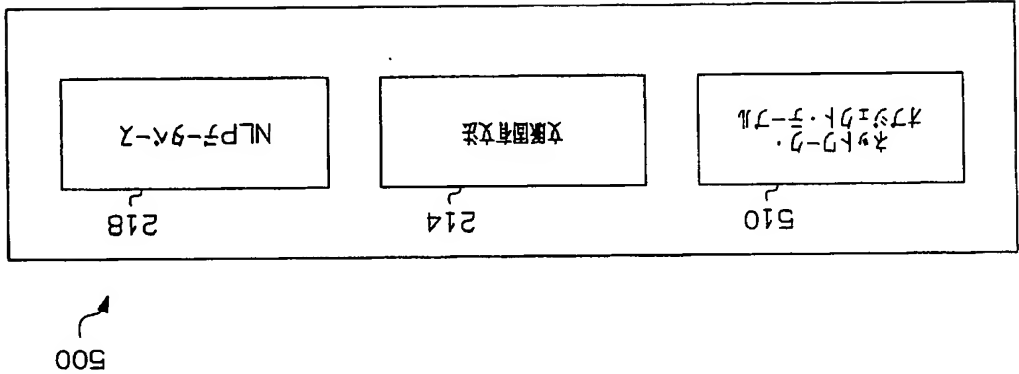
【 図 3 D 】



【 図 3 C 】



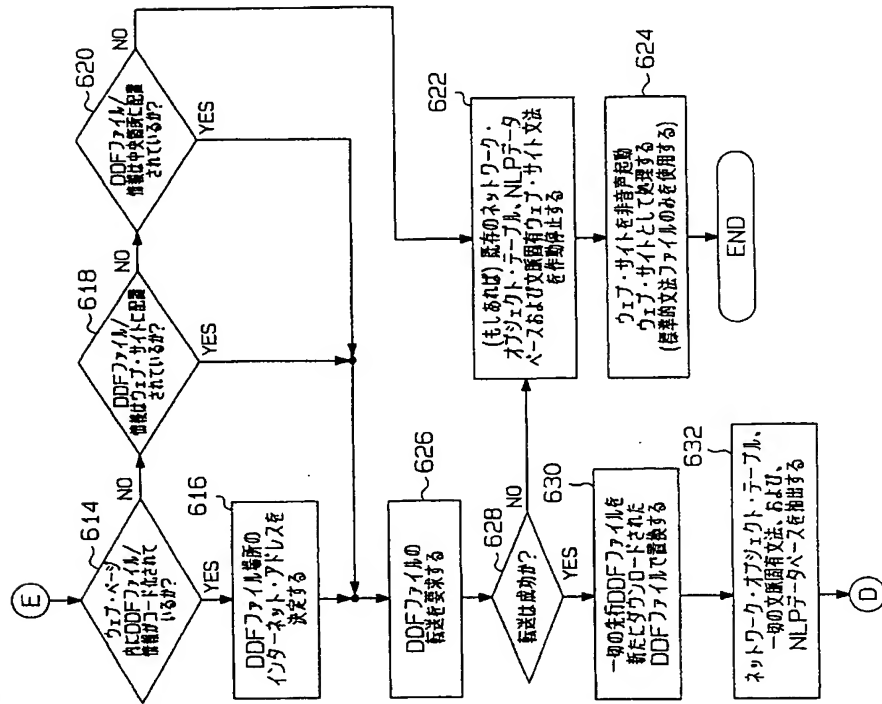
【図5】



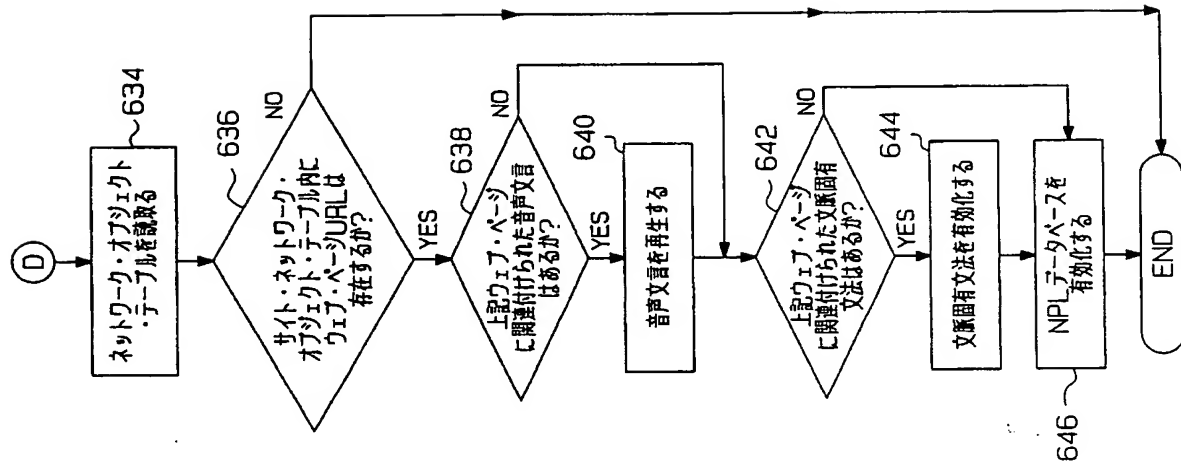
【図4】

218	400	402	404	408	410
412A	時間 (\$) にはどの映画が放映されるか (WHAT MOVIES ARE PLAYING AT \$TIME)	3	映画 (MOVIES)	動作1 映画のプレイ・サイト にプレイ入する	該当せず
412B	日付 (\$) のIBM株の値段はいくらか (WHAT IS THE PRICE OF IBM STOCK ON \$DATA)	4	株 (STOCKS)	株取引のプレイ・サイト にプレイ入する	該当せず
412C	IBMの株を (\$) F.Lで売る (SELL IBM STOCK AT \$DOLLARS)	4	株 (STOCKS)	株取引のプレイ・サイト にプレイ入する	該当せず
412D	場所 (\$) における気象はどうか (WHAT IS THE WEATHER IN \$LOCATION)	3	気象 (WEATHER)	気象のプレイ・サイト にプレイ入する	該当せず
412E	いま何時 (WHAT TIME IS IT)	2	時刻 (TIME)	時刻のプレイ・サイト にプレイ入する	該当せず
412N	ニュースを見せて下さい (SHOW ME THE NEWS)	2	ニュース (NEWS)	ニュースのプレイ・サイト にプレイ入する	イメージ を表示

【図 7 B】



【図 7 C】



INTERNATIONAL SEARCH REPORT

Information on patent family members			Serial Application No. PCT/US 99/20447	
Patent document used in search report	Publication date	Patent family number(s)	Publication date	
EP 0854418 A	22-07-1998	JP 10240493 A	11-09-1998	
EP 0747881 A	11-12-1996	US 5890123 A	30-03-1999	
		CA 2177865 A	06-12-1996	
		JP 8335160 A	17-12-1996	
EP 0854417 A	22-07-1998	JP 10275162 A	13-10-1998	
DE 4440598 C	23-05-1996	NONE		
WO 9835491 A	13-08-1998	AU 5674398 A	26-08-1998	
		EP 0958092 A	24-11-1999	

From PCT/US/2002/020447 (examined) (July 1999)

INTERNATIONAL SEARCH REPORT

Serial Application No. PCT/US 99/20447	
Documents considered to be relevant	
Category * (Check if document with indication where appropriate, if the relevant passages)	Relevant to claim No.
A EP 0 854 417 A (TEXAS INSTRUMENTS INC) 22 July 1998 (1998-07-22) column 3, line 36 - column 4, line 2 column 5, line 1 - line 14 column 5, line 29 - line 37	1-14
A DE 44 40 598 C (SIEMENS AG) 23 May 1996 (1996-05-23) abstract column 2, line 46 - column 3, line 17	1, 2, 7-9
A WO 96 35491 A (JOHNSTON ROBERT DENIS BRITISH TELECOM GB) 13 August 1996 (1996-08-13) column 4, line 28 - column 5, line 8	1, 2, 7-9

From PCT/US/2002/020447 (examined) (July 1999)

フロントページの続き

(81)指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW

【要約の続き】

習する。